REPORT RESUMES

ED 010 290 24
DEVELOPMENT OF CONCEPT FORMATION IN CHILDREN.
BY- AMSTER, HARRIETT
UNIVERSITY OF CALIFORNIA, BERKELEY CAMPUS
REPORT NUMBER CRP-2243 PUB DATE 66
REPORT NUMBER BR-5-0535
EDRS PRICE MF-$0.18 HC-$4.64 116P.

DESCRIPTORS- *CONCEPT FORMATION, *CONCEPT TEACHING, *ASSOCIATIVE LEARNING, ELEMENTARY SCHOOL STUDENTS, *COGNITIVE PROCESSES, *TASK PERFORMANCE, BERKELEY, CALIFORNIA

THE GENERAL PROBLEM UNDER INVESTIGATION CONCERNED TWO PROCESSES OF CONCEPT FORMATION - THE DEDUCTIVE REASONING PROCESS AND THE ASSOCIATIVE PROCESS - EMPLOYED BY CHILDREN OF ELEMENTARY-SCHOOL AGE. IT WAS ASSUMED THAT BOTH OF THESE PROCESSES ARE FOUND IN VARYING DEGREES AMONG CHILDREN OF DIFFERENT AGES AND MENTAL ABILITIES, DEPENDING UPON THE CHARACTERISTICS OF THE PARTICULAR TASK. THE FIRST TWO EXPERIMENTS WERE CONCERNED WITH THE INTERACTION OF BOTH OF THESE PROCESSES IN A CONCEPT FORMATION TASK, WHILE SUBSEQUENT EXPERIMENTS REPRESENTED AN ATTEMPT TO FOCUS UPON THE ASSOCIATIVE PROCESS BY STUDYING ITS OPERATION IN A VARIETY OF SITUATIONS. CONSISTENT INDICATIONS APPEARED THAT, WHEN LARGE AND SMALL VARIETIES ARE EQUAL IN STRENGTH, A SMALL VARIETY IS LIKELY TO PROMOTE CONCEPT ATTAINMENT TO A GREATER EXTENT THAN A LARGE VARIETY. IT WAS ALSO FOUND THAT IT IS POSSIBLE TO CONSIDER THE ROLE OF VARIETY IN THE ATTAINMENT OF CONCEPTS AS A FUNCTION OF THE DEDUCTIVE AND ASSOCIATIVE PROCESSES WHICH CAN BE ASSUMED TO OCCUR, ALTHOUGH IN VARYING DEGREES, AS A FUNCTION OF THE DEMANDS OF THE TASK AND SUBJECT CHARACTERISTICS. (GD)

ED010250

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Development of Concept Formation in Children

Cooperative Research Project No. 2243

Harriett Amster

University of California
Berkeley, California

1966

This research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education and Welfare.

ACKNOWLEDGMENTS

Without the full cooperation of many school districts in the vicinity of the University of California at Berkeley, this research would not have been possible. I gratefully acknowledge the extensive participation of faculty, students, and staff in Berkeley, Oakland, San Leandro, and San Lorenzo, California. In particular, I should like to thank for strong administrative support Dr. N. V. Sullivan, Superintendent and Dr. D. Freudenthal, Director of Research, of the Berkeley Unified School District; S.S. Phillips, Superintendent, M. Elliott, Director of Research, A. Tudyman, Director of Special Education, and M. Groelle, Supervisor of Classes for the Mentally Retarded of the Oakland Public Schools, A. Smith, Assistant Superintendent of the San Leandro Unified School District, and W. Dresbach, Director of Instruction, San Lorenzo Unified School District. Moreover, I am especially grateful to the many principals who devoted much effort to finding space where there was none, who rearranged schedules to make the research possible, and who gave willingly of their time in other ways. These included J. P. Matlin and W. Woodworth (Berkeley), H. D. Herrington, H. A. Blethen, L. R. Solari (Oakland), L. McCormick, M. Danielson, E. Lee, L. Wheeler, B. Miller (San Leandro), B. Miner, I. Jenkins, and O. Weston (San Lorenzo).

In addition, many teachers and countless children contributed unstintingly of their time and effort. I suppose I am most indebted to the participating children who were exceedingly cooperative and often patient in their attempts to master materials which may have seemed somewhat foreign to them.

Some of this research as noted in the text was carried out in collaboration with Douglas Chalmers and Leonard Marascuilo. One extensive study was entirely conducted and analyzed in a preliminary way by David A. Wicklund. Computer analysis of the data of most studies was supervised by Alice M. Gordon. Research assistants who conducted sizeable portions of the research included Eileen Eaton, Clark Sharick, Tom Shuell, Carol Stoltz, Karen Tankerslee, Floyd Turner, Judith Wood. The IBM 7093 at the Computer Center of the University of California at Berkeley was employed in processing the data.

Preface

In the original proposal one sample experiment was described which investigated each of five specific problems. These five major experiments were conducted to bear not only on the specific problem under which they are classified, but also on some of the other major problems. Thus, more information is available concerning each problem area than could ever be obtained from one isolated experiment or a series of experiments using only sample of Ss or one task or set of procedures. Consequently, it will be fruitful to present the experiments or integrated experimental series separately, as entities which deserve consideration in their own right and subsequently, to discuss the relevant problem areas and research objectives, crosscutting specific experiments and integrating the findings in the hope of illuminating our knowledge of children's concept learning.

The introductory chapter contains some discussion of the variables with which individual experiments deal and cogent ways in which they interlock and interrelate. The seven succeeding chapters each contain detailed descriptions of specific experiments or groups of related experiments, together with the discussion of hypotheses and research related specifically to them.

Certain chapters of this report refer specifically to the five major problem areas which were presented in the original contract, but the order in which they are presented herein has been altered. Chapter 2 deals with The Effect of Set on Concept Formation: Nonverbal Concept; Chapter 3 deals with The Relation Between Intellectual Level and Stimulus Factors; Chapters 4 and 5 deal with The Effect of Variety as a Function of Task Difficulty; Chapters 6, 7 and 8 deal with The Learning Ability of Children as a Function of Type of Mediating Relationship, I and II.

CONTENTS

Page Numbers

Acknowledgments ii

Preface iii

Tables iv

Figures viii

Chapter 1: Summation and Interference as Processes of Concept Attainment: The Theoretical Problem and the Experimental Objectives 1

Chapter 2: The Effect of Instructional Set and Variety of Instances on Concept Formation: Non Verbal Concept 5

Chapter 3: The Effect of Variety and Intelligence on Children's Concept Learning 19

Chapter 4: The Effect of Variety as a Function of Task Difficulty in the Acquisition of Word Meaning 28

Chapter 5: The Effect of Variance and Variety in Convergent Association 36

Chapter 6: The Effect of Convergent Associative Strength and Stimulus Variance on Conceptual Grouping and Learning Difficulty 42

Chapter 7: Learning Ability of Children as a Function of Type of Mediating Relationship 59

Chapter 8: Learning Ability of Normal and Retarded Children as a Function of Set 69

Chapter 9: Summary, Conclusions, and Implications 74

Appendices 81

References 103

TABLES

Page Numbers

Chapter 2

Analysis of Variance of Number of Correct Responses for the Last 24 Trials of the Pretraining Task 13

Mean Number of Correct Responses During Pretraining, Learning, and Generalization for the Major Sources of Variation 13

Mean Number of Correct Responses of the Concept-Learning Task for the Variety by Set Interaction 14

Chapter 3

Analyses of Variance for Mental Age, Learning Scores, and Test Scores 21

Average Mental Ages in Years by Grade and SES 22

Average Learning Scores by Grade and SES 23

Average Learning Scores by Form by Variety by Grade 23

Mean Scores on the Test as a Function of Grade and SES 25

Chapter 4

Median Relative Frequency of Correct Responses in the Grouped Conditions Compared with Values Obtained from the Responses to Single Sentences 31

Frequency of Correct Responses (%) According to Strict and Lenient Criteria and Expected Frequencies (%) Based on the "Best Single" and Mean Dominance for the Strict Definition of the Expected Values 33

Mean Number of Correct Identifications of Easy and Difficult Concepts for Each Condition of Variety: Lenient Scoring and Lenient Definition of Difficulty 33

Mean Number Correct (%) of Easy and Difficult Concepts for Which (Strict) and Lenient Scores are Presented for Strict and Lenient Definition of Easy and Difficult 34

Chapter 5

Mean Convergent Frequencies (%) for Adults and Children, of Responses Varying in Response Dominance (%) 39

Sets of Pairs, Triplets, and Quartets Which are Matched for Mean Dominance and Mean Variance and Specified with Respect to Part of Speech 40

Chapter 6

Mean Values (%) for Each Subgroup Within Each List of the Manipulated and other Stimulus Characteristics, Showing also the Mean Number of Correct Responses During Learning 50

Mean Number of Correct Responses, Overt Errors, and Grouping Scores in the Four Lists 51

Tables, continued

Page Numbers

The Mean Number of Overt Errors Given to High and Low Variance Doublets for Each Block of Two Trials 52

The Effect of the Two Types of Associative Strength on Number of Correct Responses, and Overt Errors 52

Mean Number of Correct Responses in Each Block of Two Trials for Doublets of Varying Associative Strength 52

Mean Number of Overt Errors Over Blocks of Two Trials as a Function of Mean Associative Dominance 53

Number of Correct Responses Within Blocks of Two Trials as a Function of the Two Types of Associative Strength and Strength of the Convergent Primary 53

Number of Overt Errors Within Blocks of Two Trials as a Function of the Two Types of Associative Strength and Strength of the Convergent Primary 53

Mean Number of Overt Errors per Block of Two Trials Given to Doublets Varying in Two Types of Associative Strength and Variance 54

Mean Number of Correct Responses in Each Block of Two Trials Given to Doublets Varying in Two Types of Associative Strength and Variance 54

The Effect on the Mean Number of Overt Errors per Block of Two Trials of the Strength of the Convergent Primary on the Relationship Between Convergent Associative Strength and Variance 54

Mean Clustering Scores as a Function of the Major Experimental Variables 55

Total Frequency of Each Type of Error Within Each List 55

The Frequency of G-Errors (%) in the Four Lists, Showing the Main Effects of Strength of Primary and Variance 55

Chapter 7

Experimental Paradigm for Mediation Experiment I, Showing Experimental Design and the Conditions as Experienced 63

Stimuli Employed in Mediation Experiment I 64

The Interaction of Response Constancy and Trials with Replication: Number Correct by Blocks of Trials, for Transfer Test 65

The Differences in Grade Levels as a Function of Replication: Trials to Criterion on List I and Number Correct on List II 65

Chapter 8

Stimuli Employed in the Second Experiment on Paired-Associate Learning in Normal and Retarded Children 72

The Mean Number of Trials to Criterion During Original Learning and Mean Number Correct on the First Transfer Trial as a Function of the List and the Replication Group 72

Tables, continued

Page Numbers

The Main Effects of the Transfer Conditions with Respect to the Mean Number of Correct Responses per Block of Two Trials and Including Comparable Results for Experiment 1 73

Number of Ss (%) Showing Above-Chance Grouping on the Transfer Task for Experiment 2 73

Chapter 9

The Effect of Variety on Deductive and Associative Concept Attainment 74

APPENDICES

Appendix 4

Materials for Experiments 2, 3, and 4 and Single Sentence Norms; Sentences Grouped by Strict Definition of Each Concept 81

Expected Values for Each Sentence for the Strict Criterion, Consisting of the Percentage of Responses to the Single Sentences Which were Considered to be Correct Responses 86

The Easy and Difficult Concepts Employed for Each Variety Condition, Presented in the Order of the Frequency with Which Responses Correct by a Lenient Criterion were Given (n = 36) 89

Total Number of Correct Responses for Each Concept Under Each Variety Condition: Lenient and (Strict) Scoring 89

Appendix 5

Triplets and Quartets Grouped by Part of Speech of Stimuli and the Convergent Response and Ranked in Mean Dominance (%) 90

Pairs Grouped by the Part of Speech of the Stimuli and the Convergent Response and Ranked Within Groups in Mean Dominance (%) 94

The Frequency of the Most Common Convergent Response in Pairs, Triplets and Quartets Matched Item for Item in Mean Dominance and within Broader Limits for the Dominance of the Best Single Word in the Set 98

Appendix 6

The Four Lists Employed in Paired-Associate Learning 99

Appendix 7

Distribution of Ss in Schools by Grade and Replication. Number of Different Classes (and Teachers) are Given in Parentheses 100

Means and Ranges of CA and IQ as a Function of Grade and Replication. Number of Ss Upon Which the Means are Based are Given in Parentheses 100

Subject Selection 101

Appendix 8

Subject Characteristics 101

Pictorial Stimuli as Paired in Experiment 2 102

FIGURES

Page Numbers

Chapter 2

Sample Problems 9

Chapter 3

Cumulative Average of Learning Scores by Number of Problems. (Shown by blocks of three) 24

Cumulative Average Transfer Scores by Number of Problems. (Shown by blocks of four) 26

Chapter 6

The Mean Number of Overt Errors Given to High and Low Variance Doublets for Each Block of Two Trials 56

Mean Number of Correct Responses in Each Block of Two Trials for Doublets of Varying Associative Strength 56

Mean Number of Overt Errors over Blocks of Two Trials as a Function of Strength of Mean Associative Dominance 57

The Mean Number of Overt Errors over Blocks of Two Trials as a Function of the Strength of Convergent Primary and Mean Associative Dominance 57

The Mean Number of Overt Errors over Blocks of Two Trials as a Function of Strength of Convergent Primary and Convergent Associative Strength 58

Chapter 7

Number of Correct Responses on the Transfer Task over Blocks of Two Trials for each Group: Experiment 1 66

Number of Correct Responses by Blocks of Two Trials for the Three Conditions of Stimulus Grouping: Experiment 1 67

Number of Correct Responses by Blocks of Two Trials for each Condition of Response Constancy: Experiment 1 68

Chapter 1

Summation and Interference in Concept Attainment: The Theoretical Problem and the Experimental Objectives

The general problem under investigation in the present series of studies concerns the processes of concept formation employed by children of elementary school age. The two types of processes which are postulated are a deductive reasoning process which involves systematic testing of hypotheses and an associative process by which concepts are arrived at on the basis of associations to the presented exemplars of a concept. It is believed that both of these processes are found in varying degrees among children of different ages and mental abilities, depending upon the characteristics of the particular task. The first two experiments concern the interaction of both of these processes in a concept formation task and the subsequent experiments represent an attempt to focus upon the associative process by studying its operation in a variety of situations ranging from a miniature concept formation task (convergent association) to paired-associate learning.

Developmental Differences

Developmental differences could occur in two ways. First, with respect to the proficiency in the use of and the tendency to employ deductive reasoning; second, with respect to the manner of operation of associative processes. With regard to the former, much evidence is being amassed in support of this hypothesis. Inhelder and Piaget (1958) have described the achievements of children in the logical solution of problems. Directly relevant to the present investigation was Piaget's description of intuitive thought in children aged four to seven as contrasted with the operational thinking of young adolescents. Earlier, Vinacke (1950) proposed similar contrasting processes of concept formation in children of increasing age and his notions, together with Piaget's formulations, provided the basis for our hypothesis that deductive reasoning would appear mainly in older children, while a basic associative process would occur in children of all ages. Osler and Trautman (1961) have found supporting evidence for this hypothesis. It was hoped that the concept formation studies described in chapters 2 and 3 would shed more light on this question.

The second possibility that the basic associative process differs in children of different ages is strongly suggested by the work of the Kendlers and their associates (1962). They have concluded in their earlier studies that the higher mental processes of nursery school children are non-mediational while the tendency to mediate increases with age. The technique which they have used throughout a large group of studies concerns children's propensity for and ability to carry out a reversal shift. This method entailed indirect measurement of the tendency to group stimuli on a conceptual basis during learning and then to transfer the mediational response in a new situation. We have attempted to devise a situation which would permit more direct measurement of the tendency to group stimuli during learning, and a variation in transfer tests which would reflect the tendency to transfer conceptual groups as a function of the conditions which obtain during transfer. These experiments which are described in chapters 6, 7, and 8 involve the learning of paired-associate doublets. These are paired-associates, S-R pairs, for which there are two stimuli for each response, e.g., fingers - hand and foot - hand. What is studied is the tendency to learn the two pairs within a doublet together rather than as isolated units, and also, inferred, is the tendency to link the two stimuli conceptually. In the Kendler's hypothesis is correct, the tendency to form conceptual groups during learning should increase as a function of mental ability but not necessarily as a function of years of exposure to language and conceptual groupings. Therefore, both normal and retardates have been studied in this connection.

The experiments in the present report deal not only with the spontaneous tendency for Ss varying in mental ability to form conceptual groups during learning, but also with their ability and propensity to employ such groups when grouping-during-learning is forced upon them by an experimental manipulation. One possible interpretation of the Kendlers' results was that the reversal shift task required the verbalization of an antonym and there is strong evidence that the frequency of antonymic associations is lower for young children than for older ones (Ervin, 1961). This suggests that tasks involving other sorts of verbal mediation could be carried out by children who have acquired proficiency in the prerequisite type of association, and the tasks we have used in chapters 6, 7, and 8, meet this requirement. But this interpretation of the Kendlers' work may be viewed as part of a more inclusive associative interpretation of concept formation. This interpretation entails the notion that concept formation occurs by the elicitation of associates to successive instances with summation or emergence of the associations common to the instances of the set, (Amster, 1965).

Other developmental hypotheses relevant to the associative process concern the number, availability, and strength of the associations characteristic of children of different ages. Moreover, differences with respect to the strength of the summative tendency or the susceptibility to interference might be postulated. On the basis of previous work, it is reasonable to infer that the number of different associations decreases with age (Palermo, 1963; Amster and Keppel, 1966), and the strength of the primary and other very common responses increases in strength (Palermo, 1963; Amster and Keppel, 1966). On the other hand, there is some evidence suggesting that young children are less susceptible to interference from extraexperimental associations than older children (Amster and Keppel, 1965). These characteristics may be evident in children's concept learning to the extent that this type of behavior depends on the associations available and the manner in which they interact. Moreover, the fact that the number of different associations decreases with age does not mean that the number of associations which can be produced by a given S also decreases with age. In fact, the converse seems to be true. Undoubtedly, bright children and older ones can elicit more responses than younger ones when this is appropriate to the task (Osler and Trautman, 1961).

Variety and How it Affects Concept Attainment Under the Two Processes

The major experimental variable which was studied throughout the project was the effect of the variety of different instances employed for concept attainment. The variable was operationally defined as the number of different instances employed, and was selected because of its relevance to the modes of operation of the deductive and associative processes. The other variables of Difficulty and Mental Ability were studied in relation to it, and Variance was investigated as a way of studying variety.

To the extent that concept attainment is a deductive reasoning process, a large variety should facilitate attainment more than a small variety since there is *a priori* knowledge that it permits rejecting any false hypotheses in fewer trials than does a small variety (Podell, 1958). Empirical evidence for this hypothesis also exists (Podell, 1958). The basis for this hypothesis hinges upon the fact that if an instance is repeated, any false hypothesis which happened to be true for that instance would be confirmed on both occasions and could not be rejected, but occurrence of a different instance allows a false hypothesis of this type to be rejected. However, there is also reason to believe that a large variety might be more difficult than a small variety because more different false hypotheses could arise on the basis of a large variety of instances than on the basis of a small variety. A similar hypothesis has been raised by Osler and Trautman (1961). Thus, particular hypotheses might be more easily rejected, but there would be more hypotheses which required rejection if the deductive process were being employed. Also, it should be mentioned that the hypotheses are believed to arise as associations to instances and are thus strictly speaking, not a function of the operation of deductive reasoning, *per se*. Consequently, insofar as concept attainment is dependent both on deduction and on the associative elicitation of hypotheses on the basis of instances, variety is a two edged sword. On the one hand, a large variety should suggest more different associations (hypotheses) which would increase the pool of available associations which might make the correct association more likely by some ways of estimating a chance basis for elicitation of a correct (particular) association. But on the other hand, there would be a greater number of hypotheses to be tested in a large variety. The superiority of the large variety over the small variety with respect to efficiency of rejection seems uncontrovertible. Consequently, the most controversial questions concern the operation of associative rather than deductive processes *per se*.

Another source of differential effectiveness of a large and small variety when deductive behavior is elicited, concerns the possible differential effectiveness of or variance among the instances within the two types of sets. Assuming that each instance is equally effective in its individual probability of eliciting correct responses, there would be no reason to consider this factor. This assumption may be made for the task employed in experiments reported in chapters 1 and 2, but for later studies, this assumption is clearly false and was in fact, the object of study. In these cases in which the instances are known to be unequal in their tendency to elicit the correct solution, the large variety would, on the average, include more "good" instances i.e., those which have a high probability of eliciting the correct concept. This would thereby raise the probability that the correct response will occur to the set as a whole. But it also raises the probability that a "poor" instance would be included which would have a very low probability of eliciting the correct hypothesis and might have only remote relevance to it. These instances could lead to false rejection of correct hypotheses, thus counteracting the facilitative effect of the "good" example. Needless to say, the exact outcomes, should depend on the particular strengths of the tendencies for the good and poor instances to elicit the correct response. If the poor instances are not very remote, they would not tend to lead to rejection, and if the good instances are not very strong in their tendency to elicit the correct response, they might not have a strongly facilitative effect. Consequently, parametric studies of this problem are called for, and attempted in the studies reported in chapters 4, 5, and 6.

Many factors must be considered in evaluating the hypothesized relative effectiveness of a large or small variety from an associative point of view. On the basis of a simple scanning model which was implied in the discussion above which concerned the interaction of deductive and associative processes, it might be said rather that a small variety should be more beneficial because fewer associations would have to be considered before the correct one occurred. This conclusion is based on the assumption that an individual would check each associate which occurred and could tell whether an association was common (whether it had occurred before). It is also assumed that he is searching for one that he recognizes. Considering an associative interaction model, by which only a limited number of associates arise and the strongest are produced overtly, the same conclusions would be drawn on the basis that interference from competing associations would be stronger in the case of the large variety in which more different competing associations would exist. But the situation is complex.

It is of some interest that the relative merits of a large and small variety has been discussed in terms of the number of different associates elicited and based on this, the probability of occurrence of the correct response and the possibility of interference from the incorrect responses. But no mention has been made of the possibility of summation or enhancement of certain specific weak tendencies to respond by virtue of repeated presentation of the same stimuli having weak tendencies as compared with presentation of different stimuli having weak tendencies. Theoretically, the summation could occur equally in both cases, but would be differentially revealed as a function of the interference in the two cases; the question is, would more interference derive from many different weak tendencies (large variety), or from relatively few rather strong competing responses (small variety). Considering the complexity of the relationship between deductive and associative behavior, it is not surprising to discover that the effect of the variety of instances has been found to differ in seemingly inconsistent ways when the work of various experimenters is compared [Adams, 1954; Callentine and Warren, 1955; Fields, 1932; Harlow, 1951; Hovland and Morrisett, 1959; Osler and Trautman, 1961; Podell, 1958; Podell, 1963 (a) and (b)]. There is reason to attribute the differences among investigators in the results obtained to enormous differences in the tasks which were employed and in the subject population which was studied. Many of these experiments have been discussed elsewhere, (Amster, 1965).

Innovations in Method

Many variables should be controlled in the study of the number of different instances presented during concept attainment and many of these may not have been controlled in previous studies or were not relevant. However, a number of novel controls were introduced in order to make more precise measurements of the effects of this variable than were possible on the basis of previous experimentation. For example, in the studies reported in chapters 1, 2, and 4, the relative frequency of particular instances was equal for the large and small variety conditions. Thus, in a large variety set of instances, each S received all the instances. In a small variety each S received only some of the possible instances (though these would be repeated in order that he had the same number of trials of exposure as Ss receiving a large variety). In the latter case, every instance would occur an equal number of times over all the Ss within the small variety. In the case of studies in which the instances could not be considered equal in difficulty, the actual difficulty was scaled before the study of the combined instances was carried out, and this was undertaken for the experiments reported in chapter 4, 5, and 6. Further, the effect of the amount of practice on particular instances was carefully controlled in all cases, and studied in the case of the last experiment reported in chapter 4.

Chapter 2

The Effect of Instructional Set and Variety of Instances on Concept Formation: Non Verbal Concept

The experiment which follows[1] concerns the effect of a particular type of unintentional set, an aesthetic set, on level of learning during pretraining and on subsequent concept formation employing the forms for which the pretraining had occurred. It also deals with the relative efficiency of a large and small variety of instances during the concept learning phase of the experiment. It was expected that developmental trends among young children would be evidenced in line with the hypotheses discussed in the following report. The study was conducted with first graders in addition to the fourth graders whose results are reported, in an effort to observe such developmental changes in the effect of the major variables. Unfortunately, we did not succeed in obtaining reliable data with six-year-olds nor even in communicating the instructions to the groups of children, indicating to us that individual administration was necessary for obtaining information with this type of task and for children of this age level.

---

[1] Reprinted from Journal of Experimental Child Psychology, Vol. 2, No. 2, June 1965
Copyright, 1965 by Academic Press Inc. Printed in U.S.A.

6

JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY 2, 192–204 (1965)

Effect of Type of Pretraining and Variety of Instances on Children's Concept Learning

HARRIETT AMSTER AND LEONARD MARASCUILO[1]

University of California, Berkeley

The effect of instructional set on pretraining performance and subsequent learning of the mathematical concepts of set-union and set-intersection was studied in children approximately ten-years-old. Verbal pretraining was found to facilitate learning during pretraining, compared with aesthetic pretraining. However, in the subsequent concept learning task, for which the same materials were employed, the *S*s who had aesthetic pretraining acquired the concepts more readily than those who had rote-learning instructions during pretraining. This difference due to pretraining was only statistically reliable among relatively low socio-economic status (SES) *S*s in the condition that received a small variety of instances of the concept. The variety of instances presented during the concept learning task did not significantly affect concept acquisition *per se*, but generalization to new instances was significantly greater when the concept had been learned originally from a small rather than a large variety of instances.

The hypothesis of Whorf (1956) provided the basis for a study by Rasmussen and Archer (1961). According to this hypothesis, language determines the categories into which we place objects and provides the labels by which discrimination among them occurs. Rasmussen and Archer (1961) compared the effects of two types of familiarization with randomly generated forms, *language pretraining,* which entailed learning a verbal label, and *aesthetic pretraining,* which entailed making aesthetic judgments. Surprisingly, the aesthetic pretraining facilitated performance in a subsequent concept identification task. The same forms were employed in both tasks, and the facilitation under aesthetic pretraining occurred when the shape of the form was the relevant dimension. To explain these results, Rasmussen and Archer (1961) inferred that *S*s making aesthetic judgments must attend to various dimensions of the stimulus forms in contrast to the others who probably responded to more limited aspects of the shape. In other words, those who learned the verbal label may have merely learned the minimum basis for discriminating between the shapes, while

---

[1] This research was supported by U.S.O.E. Project No. 2243 and was conducted at the Institute of Human Learning which is supported by a grant from the National Science Foundation.

those who made aesthetic judgments may have noticed a multiplicity of aspects.

The first problem of the present study deals with the possibility that differences between the types of pretraining depend on the pre-experimental linguistic habits of the *Ss*. Accordingly, it seemed reasonable to suppose that the relative advantage of the aesthetic pretraining might be restricted to adult *Ss*, as used by Rasmussen and Archer (1961), who probably verbalized the various dimensions of each form. It might, however, be expected that children would be less likely to verbalize or even notice such details. Rather, it would seem that an aesthetic set would, for children, involve a relatively global response to forms. Furthermore, the effect of verbal labeling might be more facilitative for children than for adults because adults would have a stronger tendency to label objects spontaneously, and therefore, not benefit from the specific instruction to verbalize to the extent that children would.

Because of the contrasting nature of the two types of pretraining, the study of Rasmussen and Archer (1961) did not include a measure of degree of learning achieved during pretraining. In view of the hypothesized differences between the instructional sets it would be of interest to know the relative levels of familiarization with the spatial forms that were achieved under the two conditions. A modification of the original experimental design was devised in order to permit independent assessment of level of learning during pretraining and subsequent level of concept learning.

On the basis of the result by Rasmussen and Archer (1958), stimuli that are familiarized under aesthetic instructions would be expected to be more discriminable than those familiarized by verbal labeling. Accordingly, any differences between instructional sets should be greater for unfamiliar forms than for familiar ones, which are presumed to be differentiated already. Furthermore, learning a nonsense-name might have little effect on the ability to identify the familiar forms because names are assumed to be available. It should, however, have a facilitative effect upon learning based on unfamiliar forms.

A second problem concerns the number of different instances (variety) that is presented to the *S* as the stimuli from which a concept is formed. Typically, large and small variety conditions consist of equal numbers of problems on the basis of concepts to be formed; but in a small variety certain instances are repeated more often than in a large variety. There are more different problems or instances in a large variety condition than in a small variety. Although acquisition of concepts under intentional set and generalization of concepts to new instances have been found to be facilitated by a large variety (Harlow, 1951; Podell, 1958), there is also some relevant contradictory evidence on this point (Podell, 1963a,b). The advantage of a large variety of instances would seem to derive from the fact that it makes possible greater efficiency in the ability to reject a given false hypothesis than a small variety. An advantage of the small variety should occur in cases where enhanced memory for particular instances would be beneficial, since it is believed that recall of particular instances would be improved under this condition. It may also be of benefit where large numbers of false hypotheses would compete with associatively remote correct hypotheses, making their elicitation less likely than under a small variety where fewer false hypotheses would be generated. It is equally important that the small variety would ordinarily entail elicitation of fewer false hypotheses, which might *in toto* be more readily rejected than the larger set of false hypotheses generated under a large variety. The pretraining variable was also hypothesized to affect the volume of associations which would occur in response to the forms and, therefore, might interact with the variety variable.

METHOD

Subjects

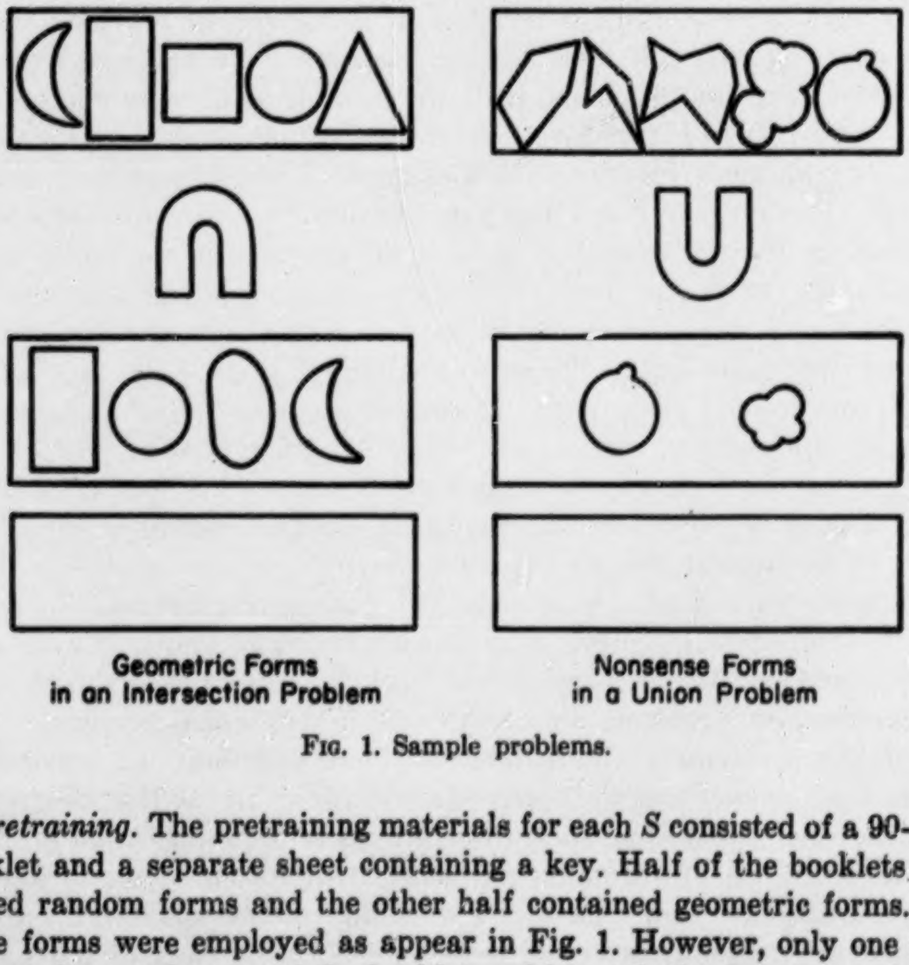
An equal number of *S*s were selected from each of two 4th-grade classes in two schools of the Berkeley Unified School District, Berkeley, California; and they were assigned pre-experimentally to the treatment groups by randomized blocks on the basis of MA. Only 47 out of the 48 *S*s were actually present for the experiment. Lorge-Thorndike IQs were available for 41 of the *S*s and MAs of the others were estimated on the basis of performance on achievement tests. The MAs ranged from 8.8 to 12.3 years with a mean of 11.0 years and a standard deviation of 1.1 years; the CAs ranged from 8.5 to 12.1 years with a mean of 9.9 years and a standard deviation of 0.4 years. Children having difficulty with English were excluded. Procedures were administered in entire classes. One of the classes was high in socio-economic status (SES), and the other was about average.

Experimental Design

The experiment consisted of two main parts, pretraining and concept learning. The same *S*s participated in both parts; and otherwise all treatments were administered to independent groups. A two-by-two orthogonal design was employed in which the groups received *random* or *geometric* forms and *verbal* or *aesthetic* instructions. For the learning task the *S*s received one of two types of variety of different instances of the concepts to-be-learned, *small variety* or *large variety*. For both pretraining and learning there was an additional two-level variable that reflects not only the two classes but also two different page-sequences in which the booklets for pretraining and concept learning were arranged. In addition, different experimenters were used in each class. These variables are confounded with classes.

MATERIALS

There were six forms of each type, random, and geometric. The randomly-generated forms were constructed according to Method I of Attneave and Arnoult (1956). For the angular figures, straight lines were used to connect the sets of 5, 7, and 9 points; for the curved figures, curved lines were used to connect the set of 3, 4, and 5 points. Examples of the concept-learning problems that employ the forms appear in Fig. 1.



Geometric Forms
in an Intersection Problem

Nonsense Forms
in a Union Problem

Fig. 1. Sample problems.

Pretraining. The pretraining materials for each *S* consisted of a 90-page booklet and a separate sheet containing a key. Half of the booklets contained random forms and the other half contained geometric forms. The same forms were employed as appear in Fig. 1. However, only one form appeared on each page of the booklet and beneath the form was a box in which *S*s were to write their response. The booklets were arranged in such a way that each form appeared once in each block of six pages. There were 15 blocks of six pages, and the forms were arranged in a different random sequence in each block. There were two different sequence assignments, A and B; however, one assignment was used for each type of form in each class. Thus, there were actually four different types of pretraining booklets. One class received booklets containing both types of forms in the same page-sequence, and the other class received two types of booklets in the other sequence. This assignment of booklets enabled the same response to be correct on every page of every booklet within each class.

There were two keys, one for each set. For the aesthetic condition, the key consisted of a list, arranged in order of magnitude, of the six categories ranging from "very ugly" to "very pretty" and their corresponding numbers from one to six, respectively. For the verbal condition, it consisted of six trigrams: *paf, fub, nax, zut, bep, tev*, and their corresponding numbers arranged in order from one to six.

Concept learning. Each *S* received a booklet of 40 problems, one on each page. Of these, 36 were concept-learning problems; and the last 4 were generalization problems. Half were problems of set-union and half of set-intersection. Figure 1 consists of two sample pages, one a set-union example and one a set-intersection example. The problem was arranged the same way on each 3 × 4 inch page. Two to five forms were reproduced in a box on the top row; a large **U** or **∩** appeared in the middle of the second row; two to five forms were reproduced in a box on the third row; and the fourth row was an empty box in which the *S*s were to write the numbers corresponding to the answer. Although every *S* did not have the same problem on a given page, the correct response to each page was the same for all booklets within each class. The same type of problem, i.e., union or intersection, appeared on a given page of all booklets within a class. The only exception was that each booklet contained nine blank pages in addition to the 40 pages on which there were problems. These blank pages were inserted to give the *S*s breaks during learning.

Half of the booklets were large variety booklets, and half were small variety booklets. In the large variety booklets there were eighteen different intersection problems and eighteen different union problems. Even though the problems were all different, there were only six possible answers. Each answer was the correct answer for six of the thirty-six different problems. Thus, there were three groups of six intersection problems and three groups of six union problems, with the problems in any one group having the same answer. The small variety booklets were constructed by randomly selecting two problems from each of the six different groups. These two problems were each repeated three times in a given booklet. The problems for the small variety booklets were selected so that each of the thirty-six different problems appeared equally often in the total set of booklets. Consequently, the particular problems in any two small variety booklets would tend to differ. The sequence of problems within each booklet was randomized with the restriction that the same answer was correct on a given page for all booklets within a class.

There were six blocks of six different correct answers that occurred in a different random sequence on each repetition. Two different sequences of correct answers were employed for the two classes.

Generalization. The last four pages of each booklet contained the same intersection and union problems, two of each. Instead of being implemented by forms, one problem of each type employed the words *cow, fox, dog, cat, rat,* and *pig,* and the other employed the letters, *X, W, E, J, F, K, B,* and *H.* The order of the four pages was randomized in each booklet so that the same answer was not correct on a given page.

PROCEDURE

Pretraining. Half of the *Ss* from each class were assigned to the aesthetic set condition and half were assigned to the verbal condition. For the aesthetic instructions, *Ss* were asked to learn how much other children liked each shape. *Ss* were encouraged to guess at first and told that they would start to learn what other children thought of the forms. They were to record their guess by writing a number from one to six in response to each form. For the verbal instructions, *Ss* were asked to learn the name of each shape and to record the number from one to six corresponding to each name. *Ss* were to guess at first and told that they would start to learn the names of the forms they would see. For both sets, the instructions were paraphrased in various ways; and the children were questioned about them to insure that the task was clear. For one of the classes the aesthetic treatment was administered in the classroom while the verbal treatment was administered in another room, and for the other class the assignment of treatment to room was reversed.

Among the *Ss* receiving each instructional set, half of the *Ss* (six) were given the nonsense figures and the other half were given the geometric forms. Each *S* was given a key appropriate to his set condition and a pretraining booklet.

The *Ss* were given 15 seconds to respond to each page for the first 12 pages and thenceforth, 10 seconds per page. *E* said, "pencils down," and gave the feedback which included the nonsense name or aesthetic judgment in addition to the correct number. After a five second intertrial interval, *Ss* were told to "turn the page" and the procedure was repeated. The same number was correct in a given treatment condition despite the fact that some booklets contained geometric forms and others contained random forms. At the end of 30 and 60 pages, *Ss* were given three-minute rests during which they were allowed to stretch, move around the room, sharpen pencils but not communicate with one another.

Learning. Following the pretraining task, *Ss* went outside for a 10 minute recess. The learning task was then administered in intact classes. Each *S* was given the pre-assigned problem booklet and if misplaced, another key. He was instructed as follows:

> Please do not turn the pages of your booklet until you are told to. Do not talk about the problem or what you think the answer is. Write the answer in your booklet. On each page you will see three boxes and a sign. The sign can be like this $\cup$ or like that $\cap$; and the last box is empty. (*E* wrote $\cup$ and $\cap$ on the blackboard.) You are to fill the last box with numbers from your key. At first you will have to guess which numbers to put in, but after a while you may figure out how to do it. Let's all do the first page together. Look at the first page. Decide which shapes really go into the empty box. Sometimes it will be more than one, sometimes just one. Look at the key and see what the numbers are. Write the numbers in your box. (Pause) Stop. I will put the answer on the board. If you wrote the answer this way, *125*, it is correct; *512* or *215* would also be right.
>
> Some of the pages in your booklet will be blank. If your page is blank, just leave your booklet open to that page and rest. The blank page means rest for this problem. After this, do your own work and do not write on the key. Turn the page only when told to. The pages will be timed. After you hear the answer try to figure out why that answer was correct. Ready, turn to page 2. Pencils down. The answer is ——.

For the first 12 problems, *Ss* were permitted 17 seconds to record their response and the intertrial interval was 15 seconds. For the remaining problems they were permitted 15 seconds with a 10-second intertrial interval. The feed-back was given at the beginning of the intertrial interval. *Ss* were told to spend the intertrial interval in trying to see why the answer placed on the board was correct.

Generalization. For the generalization problems, *Ss* were instructed to try to solve the problem and as soon as they finished a problem to turn the page themselves. The time allowed for the four problems was 35 seconds.

RESULTS

Pretraining Task

Since the correlation between mental age and total score was quite small ($r = 0.18$, $p > 0.05$), mental age was not used as a covariate as originally planned. This small correlation reflected the minimal range in intelligence previously mentioned. Since the two classes came from different socio-economic strata, an analysis of variance involving a nested design was deemed appropriate (Winer, 1962).[2] All statistically signifi-

---

[2] Since the analysis of variance requires that the variances between groups should represent only chance deviations from equality, Cochran's Test of Homogeneity of Variances was made, and accordingly, the hypothesis of equal variance was not rejected. The assumption that the data are normally distributed was tested by inspection of the error histogram.

TABLE 1
Analysis of Variance of Number of Correct Responses for the Last 24 Trials of the Pretraining Task

Source	df	MS	F
Between classes	1	102.08	4.77*
Within classes	45	21.36	
Set (S)	2	92.08	4.57**
Form (F)	2	6.08	—
S × F	2	0.47	—
Block	4	14.55	—
Error	35	20.13	
Total	46	23.11	

* $p < 0.05$.
** $p < 0.025$.

cant results were further analyzed by the Scheffé Method of Multiple Contrasts. According to these methods the overall differences between classes and between sets were significant at $p < 0.05$. The mean number of correct responses under the aesthetic set was 17.9 and under the verbal set 21.3. For the high SES class the mean was 21.1 and for the lower SES class the mean was 18.2. By using the Scheffé Method of Multiple Contrasts it was found that the set differences were significant for the lower SES class and not for the high SES class. However, the differences were in the same direction and the overall difference was also significant. The means for the major variables are shown in Table 2.

TABLE 2
Mean Number of Correct Responses during Pretraining, Learning, and Generalization for the Major Sources of Variation

Source	Level	Pretraining[a]	Learning[b]	Generalization[c]
Set	Verbal	21.3	11.8	1.1
	Aesthetic	17.9	14.6	1.1
Forms	Random	19.8	11.6	0.9
	Geometric	19.4	14.8	1.3
Variety	Small	19.8	13.6	1.6
	Large	19.5	12.8	0.6

[a] Mean number correct on the final 24 trails.
[b] Mean number correct on the final 24 trials.
[c] Mean number correct out of 4 problems.

Concept Learning Task

The dependent variable in the analyses to be described was the total number of correct responses for the learning task, omitting the first 12 problems. For the combined results the correlation between the dependent variable and mental age was not statistically reliable ($r = 0.25$, $p > 0.05$), unlike the correlation with total score during pretraining ($r = 0.37$, $p < 0.01$). The corresponding correlations between the total score during pretraining and the concept learning score for the 12 union and 12 intersection problems were, respectively, 0.27 ($p > 0.05$) and 0.38 ($p < 0.05$). Analyses of variances were conducted for the union problems, the intersection problems, and both problems combined. None of the sources of variation proved to be statistically significant. However, the manipulated variables were not completely without effect. In the analysis of both problems the Set by Variety interaction ($F = 3.05$, $df = 2/27$) within the classrooms was significant at $p < 0.10$. The interaction was found statistically reliable by the Scheffé method for the group from a relatively low SES area. Table 3 presents the relevant means.

For this group, there was a nonsignificant trend for performance to be facilitated by the large variety when the verbal set was given during pretraining, but when the aesthetic set was given, performance tended ($p < 0.10$) to be facilitated by the small variety. Further, among the *Ss* trained under the Small Variety, those who had received the aesthetic instructions during pretraining were significantly superior to those who had received the verbal instructions ($p < 0.05$).

TABLE 3
MEAN NUMBER OF CORRECT RESPONSES ON THE CONCEPT-LEARNING TASK FOR THE VARIETY BY SET INTERACTION[a]

	Low SES		High SES	
	Verbal Set	Aesthetic Set	Verbal Set	Aesthetic Set
Large Variety	13.8 (21.0)	9.7 (15.3)	12.2 (21.5)	15.7 (20.0)
Small Variety	8.0 (20.7)	15.8 (15.7)	13.3 (22.2)	17.3 (20.7)

[a] Pretraining means appear in parenthesis.

The pretraining scores that appear in parentheses in Table 3 clearly indicate that performance during concept learning cannot be accounted for by level of pretraining. For the significant concept learning effect, the prior pretraining scores were in the opposite direction.

Generalization Problems

Nonparametric analyses of the results of the four generalization problems were carried out. The rank correlation coefficient between the total number correct for the last 24 trials of the learning task and the number correct on the generalization problems was found to be 0.51, significant beyond the 0.01 level. The mean number correct was evaluated as a function of experimental conditions.

The Mann-Whitney Test was used to test for the significance of the main effects and the Kruskal-Wallis to test for the significance of the interactions. The main effect means appear in Table 2. The only reliable difference ($p < 0.001$) was the main effect of variety: more generalization problems (1.63) were attained under the small variety condition than under the large variety condition (.59).

DISCUSSION

In contrast to our hypothesis that verbal pretraining should have facilitated concept-learning in children to a greater degree than aesthetic pretraining, the converse was observed. This advantage of the aesthetic pretraining, which occurred only for the small variety condition, was consistent with that for college students (Rasmussen and Archer, 1961). In their study there was no measure of the degree of learning of the forms during pretraining. In the present study however, separate achievement measures were derived for pretraining and concept learning. However, it must be admitted that the nature of the aesthetic set employed in this study is distinctly different from that employed by Rasmussen and Archer (1961); and for that reason, the results of the two studies may not be comparable. In the latter case, the *S*s made aesthetic judgments while in the present study judgments were supplied to them. Curiously enough, the results for pretraining contrasted with those for concept learning. During pretraining, verbal labeling led to more efficient assignment of numbers to spatial forms than aesthetic judgments. Consequently, Rasmussen and Archer's (1961) result concerning concept learning should not necessarily be attributed to a higher level of pretraining under the verbal set. Moreover, the important possibility that *S*s must learn different things during the two types of pretraining now becomes more firmly established.

The relative advantage of labeling during pretraining would seem to be related to several differences in the responses that occur under the two conditions. One important consideration is that *S*s' own aesthetic judgments may have been at variance with the required assignments of numerals to forms, and this would have produced interference learning. It is reasonable to assume, moreover, that the established response tendencies with respect to aesthetic judgments were stronger than any tendencies they may have had with respect to the verbal labeling. Consequently, there would have been greater interference in learning the response as an aesthetic judgment than as a verbal label. But, despite this difficulty in learning the labels, *S*s may well have learned more about the stimuli during the aesthetic pretraining, as suggested by Rasmussen and Archer (1961). More explicitly, this might have entailed learning the label in response to many weak cues especially if *Ss* constantly sought new cues. This would contrast with the *Ss* in the labeling group who probably employed fewer differentiating aspects of the stimuli as cues for the response.

The fact that the verbal pretraining task involved learning an additional response (nonsense syllable) might account for the pretraining differences because this requirement might be expected to impede learning of the numbers. It might be mentioned that for neither pretraining condition was it necessary for the *Ss* to learn anything but the association between the spatial form and the number. However, we assume that they followed their respective instructions and learned more than this association. Also, there are possible benefits or interferences under either type of pretraining that could come about by trying to associate the numbers with the number of sides or other numerical aspects of a figure.

The shift in relative advantage of the two pretraining conditions could have a definite relationship to the requirements of the concept-learning task. During pretraining the *S* merely had to learn the label for each spatial form when it appeared alone. During concept learning he had to discriminate among many forms appearing together and match them. Quite possibly, the prior learning of many aspects of each form aided the *Ss* in making the complex discrimination required; and, therefore, the aesthetic pretraining facilitated the concept learning. The advantage of the aesthetic pretraining occurred primarily for the small variety condition and was significant only for the lower SES group. Conceivably, presenting the stimuli in many different arrangements during concept learning facilitated the detection of multiple cues with which to discriminate among the spatial forms. In this case the prior learning of such cues would have been less crucial than in the Small Variety case in which the repeated presentation of the spatial forms in the same positions would have made for greater reliance on prior learning of varied cues.

An additional possibility is that retention of the number-form associations and of the discriminations learned during pretraining were facilitated by aesthetic mediators that might have been acquired during aesthetic pretraining. This aesthetic mediator was no doubt more familiar, more meaningful, more pronounceable, and of greater value than a nonsense word; and mediators having these characteristics should be more effective in mediating recall. In addition, it is also interesting to consider that recall of the response was facilitated by the occurrence of many weak cues during the learning, rather than by the occurrence of a few strong cues as would have characterized the aesthetic and rote pretraining, respectively.

Several variables were found to affect concept learning. Although the variety of instances presented during concept learning was not a significant factor during this phase of the experiment, it proved to have a significant effect on the extent to which *Ss* could generalize the concepts to new instances. Those who formed the concepts originally on the basis of the small variety of instances showed significantly greater generalization to new instances than those who formed on the basis of the large variety of instances. This result is directly opposite to that obtained in widely different situations (Harlow, 1951; Callantine and Warren, 1955; Podell, 1961 and 1963). On the other hand, a large variety has been found to produce interference in some situations (Osler and Trautman, 1961; Adams, 1954), and Podell (1963a,b) found some evidence for an advantage of a small variety over a large variety in two formally similar situations by using quite different tasks. In one case, the result occurred only for fourth graders who had relative difficulty in learning the required concepts. This group may be similar to the present low SES group for whom the advantage of the small variety was more pronounced. Since there seems to be no rational explanation for a difference in generalization without an *a priori* difference in learning, we must infer that the *Ss* did in fact acquire the concepts to a higher degree under the small variety of instances.

REFERENCES

Adams, J. A. Multiple vs. single problem training in human problem solving. *J. exp. Psychol.*, 1954, 48, 15.

Amster, H. The effect of instructional set and variety of instances in children's learning, 1964, unpublished manuscript.

Attneave, R., and Arnoult, M. D. The quantitative study of shape and pattern perception. *Psychol. Bull.*, 1956, 53, 452–471.

Callantine, Mary F., and Warren, J. M. Learning sets in human concept formation. *Psych. Rep.*, 1955, 1, 363–367.

Harlow, H. F. Thinking. In H. Nelson (Ed.). *Theoretical foundations of psychology.* New York: Van Nostrand, 1951.

Osler, Sonia F., and Trautman, Grace E. Concept attainment II—effect of stimulus complexity upon concept attainment at two levels of intelligence. *J. exp. Psychol.*, 1961, 62, 9–13.

Podell, H. A. Two processes of concept formation. *Psychol. Monogr.*, 1958, 72.

Podell, H. A. The effect of the variety of instances on the production of verbal concepts. *Amer. Psychologist*, 1963, 18, 382. (Abstract) (a)

Podell, H. A. A developmental study of recall of instances as a function of set and the variety of the instances presented. *Amer. Psychologist*, 1963, 18, 346. (Abstract) (b)

RASMUSSEN, E. A., AND ARCHER, E. J. Concept identification as a function of language pretraining and task complexity. *J. exp. Psychol.,* 1961, 61, 437–441.

WHORF, B. L. *Language, thought and reality.* New York: Wiley, 1956.

WINER, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

Chapter 3

The Effect of Variety and Intelligence on Children's Concept Learning

The effects of variety upon the learning of the mathematical concepts of Boolean set union and intersection were studied. Ss from 2nd, 5th and 7th grades were drawn from 4 schools representing high and low SES areas. They were randomly assigned to 4 experimental conditions which varied in two ways. Ss in the large variety conditions were given 48 different problems to solve. Subjects in the small variety conditions were given 8 problems to solve with each problem repeated 6 times. The stimuli employed with one-half of the Ss were familiar geometric forms for which a verbal label was readily available. Stimuli for the remaining Ss were unfamiliar random forms which they had never seen before. Data from the 2nd graders are not reported because the task was, on the whole, far too difficult for them. On later testing with different stimuli it was seen that the effect of variety was significant for the low SES Ss in both 5th and 7th grade; they performed more efficiently under the small variety conditions and with the nonsense forms.

The stimuli employed for this study were the same as those employed for the previously described study. However, no pretraining procedures were employed, in order to simplify the study of the variety variable and to enable concentration upon the relation between intelligence and concept learning. Having learned of the difficulty of teaching the concepts of set union and intersection to six-year-olds, we attempted to teach these concepts to seven and eight year olds as part of the present study. Again, we found great difficulties in the administration of the group procedures and abandoned these age groups in favor of groups which were considerably older, being approximately eleven and thirteen years of age. Also, since we wished to deal with intellectual level as a variable, we also decided to manipulate SES which would entail a different manner of manipulating intelligence. We thus studied children of two different mental ages at each chronological age level. Futhermore, it enabled the investigation of the possible distinction between social intelligence and intellectual intelligence as a variable in itself.

The Effect of Variety in Children's Concept Learning[1]

Leonard Marascuilo and Harriett Amster

It is generally assumed that the learning of a new task with an improved ability to transfer the recently acquired skill to new but similar situations is facilitated by giving subjects a large set of experiences or a variety of examples during the learning period. Gagné, Mayor, Garstens, and Paradise (1962), who used a programmed task involving the learning of mathematical concepts, reported that acquisition was not affected by the amount of repetition of examples of the concepts. However, he (Gagné, 1963) found that retention was affected by this variable. Retention of a newly acquired task after nine weeks was significantly low for subjects given a minimal variety of examples but no difference in retention was found for the groups given greater amounts of experiences. The results were somewhat puzzling, however, since Ss given no examples performed as well as those given various degrees of variety. Similarly, in a study involving a task which was highly similar to that of the present study, Amster and Marascuilo (1965) found a trend for learning of a generalizable sort to be facilitated by a small number of different instances being employed during concept acquisition. Gagné's Ss were 7th-grade students who employed the familiar number system to learn the addition of integers. Amster and Marascuilo (1965) used 5th-grade Ss and a task involving acquisition of the concepts of the union and the intersection of mathematical sets.

This research extends the study of the effects of variation upon the learning of a new concept with its transfer to different but similar situations by introducing a new task and by using constant stimuli for each S. Like Gagné's (1962, 1963) task, the task employed in this study is from the area of mathematics. More specifically, the simultaneous acquisition of the concepts of set union and set intersection were studied. Furthermore, half of the Ss were given a small variety of instances while the remaining half were given a large variety of instances. The expectation would be that the Ss given the largest variation in training would learn to differentiate between set union and set intersection problems sooner than those given a small variety of experiences and then show a greater facility in transferring the acquired knowledge to new stimuli.

---

[1] In press, California J. of Educ. Res., 1966.

Procedure

Learning Booklets

In order to control variables which might have affected the performance of the Ss in the experimental situation and in order to counterbalance potentially important stimulus variables, a balanced design which entailed a scheme for constructing booklets was devised. Only the important points in the booklet-construction will be reported here.

Four different kinds of learning booklets were constructed. These correspond to the four experimental conditions of the experiment. For this experiment Ss were expected to learn the mathematical concepts of set union and set intersection. The stimuli employed with one-half of the experimental Ss were familiar geometric forms for which a verbal label or name was readily available. The stimuli employed for the remaining half of the Ss were unfamiliar random forms which subjects had never seen before. One-half of the booklets contained 48 different problems. The remaining half of the booklets contained only 12 different problems, each of which was repeated somewhere in the booklet four different times. Thus, the four experimental conditions were 1. Familiar forms in a large variety of problems. 2. Familiar forms in a small variety of problems. 3. Unfamiliar nonsense forms in a large variety of problems. 4. Unfamiliar nonsense forms in a small variety of problems.

Each booklet consisted of eight blocks of seven pages, or 56 pages. One block consisted of three different union problems, three different intersection problems, and one blank page which served to balance the design and to give each subject a rest period in each block. The large variety booklets contained 24 different union problems and 24 different intersection problems. The small variety booklets contained six different union problems, each repeated four times, and six different intersection problems, each repeated four times. These problems were selected so that each of the 48 different problems appeared the same number of times throughout the entire experiment.

The problems were constructed so that each nonsense and geometric form appeared the same number of times within a booklet so that no particular form was favored or under-emphasized. Furthermore, the same six responses (answers) were employed equally in all conditions. Within the 24 union problems there was a further balanced subdivision of easy and hard problems. This distinction was based upon the number of figures that were used on a particular page. Easy problems employed five figures divided into two sets, one set with two figures and one set with three figures. Hard problems employed seven figures divided into two sets, one set with three figures and one set with four figures. Easy and hard problems were balanced throughout all booklets.

Subjects were given immediate feedback after each problem. Each page of the booklet had an underfold on the bottom of the sheet. On the underfold was a box exactly like the answer box used by the Ss. It differed from the one used by the Ss in that the answer was marked by means of large X's drawn over the correct forms. Being immediately below the answer box, Ss were given the opportunity to make direct and immediate comparisons with their own responses.

Test Booklets

The test booklets were identical for all subjects. They contained 36 pages with 16 union problems, 16 intersection problems, and four blank pages. The stimuli employed in the test booklets differed from those used in the learning booklets. The stimuli were of four types: letters of the alphabet, English nouns, pictures of well known objects, and flags with various designs. Each booklet consisted of four blocks of nine pages each. Each block contained four union problems, four intersection problems, and one blank page.

Subjects

Subjects were obtained from four different schools in the Berkeley Unified School District. Two of the schools represented neighborhoods of upper socio-economic status (SES) and two represented neighborhoods of lower SES. One fifth-grade and one seventh-grade class was chosen from each SES area. Within each class 16 Ss were pre-experimentally selected by random means to take part in the experiment. They were then randomly assigned to the four experimental conditions. The total sample size was therefore 64.

Procedure

Each S was given a learning booklet and wax color crayon to mark his answers in the booklets. The Ss were given the following instructions and information.

"We have some problems that we would like to have you solve. At first you will have to guess. These problems are printed on each page of the large booklet that we have given you . . . After working on a few of the problems, some of you may discover the correct way for finding the right answer . . . The booklets have been made so that not everyone is resting at the same time. While some of you are resting, others will be solving problems. When you have a rest period, sit quietly and wait for us to move on to the next page in your booklet . . . On the top of the page you will see a box in which there are some pictures or shapes. Beneath this box you will find a large U. Sometimes it will be upside down, and sometimes it will be right side up. Beneath the U, you will find another box in which some pictures or figures have been drawn. The two boxes and the U are the problem for that page. When the U points up, the problem can be solved by a certain rule that you must find. When the U points down, a different rule is needed to solve the problem. Below the problem you will find another box in which there are six figures. This is the answer box . . . After you have decided what you think is the answer to the problem, take the crayon we have given to you and draw an X over the pictures or figures that you think are the answer. After everyone has made a guess and drawn his X's over the pictures, we will all look at the right answer . . . If you should give the wrong answer, you should then study the problem boxes and find out why your answer was wrong so that you can do better on the next problems."

Subjects were given 15 sec. to solve each problem and 15 sec. to study the problem and the correct answer together. This gave them immediate feedback after each problem. After the completion of the learning task, the Ss were given the test booklets. The procedure used for the learning booklets was repeated except that the Ss were not given any feedback after the problems. Whereas for the learning task the Ss were exposed to a particular problem for 30 sec., they were exposed to the test problems for only 15 sec.

Results and Discussion

Mental Age

If performance on the task were found related to mental age, a covariance adjustment was planned. The analysis of variance for mental age is shown in the table. Of the total sum of squares, 77% is associated with the differences between the schools. The F-ratios measuring

Analyses of Variance for Mental Age,
Learning Scores, and Test Scores

Source of Variation	df	Mental Age	Learning Scores	Test Scores
		F	F	F
Between Schools	3	65.5*	32.5*	16.3*
Grades (G)	1	36.6*	13.1*	1.6
SES	1	141.9*	88.4*	47.2*
G x SES	1	17.7*	a	a

continued on next page

Analyses of Variance for Mental Age, etc., continued

Source of Variation	*df*	Mental Age *F*	Learning Scores *F*	Test Scores *F*
Within Schools (mean square)	60	(235)	(93)	(84)
Variety (V)	1	a	a	a
V x G	1	a	a	a
V x SES	1	a	3.5**	a
V x G x SES	1	1.2	a	a
Form (F)	1	a	a	a
F x G	1	a	1.6	a
F x SES	1	a	a	1.4
F x G x SES	1	a	a	1.3
F x V	1	a	a	a
F x V x G	1	a	6.1*	a
F x V x SES	1	a	a	a
F x V x G x SES	1	2.3	a	3.8**
Residual (mean square)	48	(253)	(90)	(86)
Total	63			

a $F < 1$

* $p < .05$

** $P < .10$

the differences between Grades, SES, and the Grade by SES interaction are all statistically significant. The mean mental ages for these statistically significant sources of variation are summarized in the following table. (None of these differences were unexpected.) The

Average Mental Ages in Years by Grade and SES

SES	Grade		
	Fifth	Seventh	Both
Low	9.6	10.2	9.9
High	12.0	15.3	13.6
Both	10.8	12.7	11.8

grade difference merely reflects the two year difference in age between 5th and 7th grade students. Although it is well known that the mean mental age of students in low SES schools is lower than that of students in high SES schools, the magnitude of the difference, 3 1/8 years, is noteworthy. The significant Grade by SES interaction may only be an artifact resulting from the kinds of ability groupings used by the various school administrators. However, this probably does not explain why the mean mental age for the low SES 7th grade students is less than that of the high SES 5th grade students. A serious attempt was made to employ exactly comparable 5th and 7th grades within each SES group.

For all sources of variation that involved any of the experimental conditions, mental age was uniform and all *F*-ratios are nonsignificant. Since mental age was uniform over all the experimental conditions, it was decided that statistical control of mental age by means of an analysis of covariance was not required.

Learning Scores

The correlation coefficient between mental age and the learning scores was .80, where learning scores are defined to be the total number of correct solutions or answers. With this as a criterion variable, the maximum learning score available was 48. The mean learning score was 24, the median score was 23.5, and the standard deviation was 14.5.

The analysis of variance for the learning scores is shown in the first table. Of the total sum of squares 57% is associated with the differences between the schools. This is not surprising since the differences in mental ages between the schools were significantly large. During the testing period most Ss in the low SES schools appeared to "give up" when they found that after a dozen or so trials they had not solved the problem. A very common complaint from many of them was that the problems were unsolvable. At the other extreme, the high SES Ss seemed to find the problems challenging and stimulating. Most likely these differences reflect previous experiences with problem solving. One would suspect that high SES children had successful experiences in solving problems, while the low SES children had not. As a group, the low SES Ss did not find the task stimulating and showed signs of relief when the testing was over. The average learning scores by grade and SES are shown in the following table.

Average Learning Scores by Grade and SES

SES	Grade		
	Fifth	Seventh	Both
Low	11.7	15.7	13.7
High	30.1	38.4	34.3
Both	20.9	27.1	24.0

Cumulative learning curves for the four groups of Ss are shown in Figure 1 on the following page. As can be seen, most of the high SES 7th grade Ss had learned the task immediately after the 21st problem. This decision is based upon the fact that from these points on, the cumulative learning curves are linear. The magnitude of the slope of the learning curves suggests that performance was nearly perfect following the 15th and 21st problems respectively. The same did not hold true for the low SES subjects. Their learning curves continued to accelerate, but at an extremely slow rate. This most likely reflects their "giving up" on solving the problems.

Finally, there was a significant Form by Variety by Grade interaction in the analysis of variance. According to the Scheffé method of multiple contrasts, the Variety by Form interaction was only statistically significant within the 5th grade. The means corresponding to this significant source of variation are shown in the following table. As can be seen, for the 5th grade Ss receiving the small variety condition nonsense forms were more beneficial for learning the concepts of set union and set intersection than familiar forms. (This result is contrary to expectation, and why it should hold for the 5th grade Ss is problematical.) It could be that ready names for the geometric figures and experience with standard contexts interfered with Ss' ability to use them in different situations where they are not normally applied. Why this difference between small and large variety problems of experiences exists is not certain. It could be that once the Ss in the small variety conditions recognized that the problems were repeated they began to pay more attention to the task and then learned the concepts for the problems that they began to recognize. In other words, discrimination among forms may have

Average Learning Scores by Form by Variety by Grade

Forms	Fifth Grade		Seventh Grade	
	Small Variety	Large Variety	Small Variety	Large Variety
Nonsense	27.0	17.3	22.5	23.1
Geometric	17.5	21.8	24.5	21.9

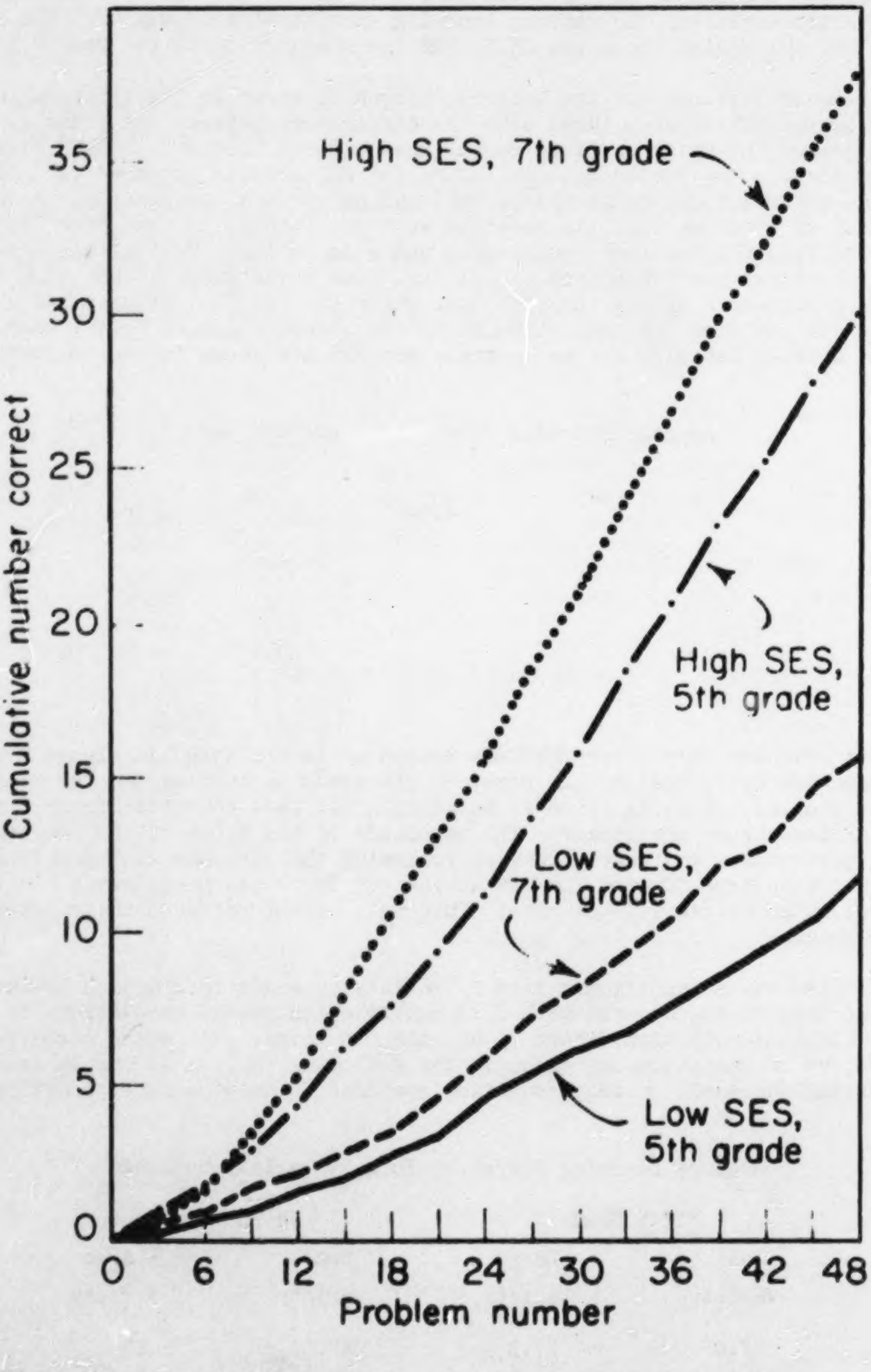


FIGURE 1. Cumulative Average of Learning Scores by Number of Problems. (Shown by blocks of three)

been more of a problem for 5th graders than for 7th graders and the small variety may have facilitated this discrimination among forms. Moreover, the facilitation by small variety appears to have affected the random forms to a greater extent than the familiar forms which may already have been differentiated.

There is some evidence that some Ss did not distinguish between the union and intersection symbol and failed to realize that two different concepts were to be learned. This hypothesis is supported, for the most part, by a post hoc analysis of the responses made by each S in the learning booklets. For this analysis, the problem immediately following each correct response was examined to determine if it was of the same type (union or intersection) as the preceding problem. If the two problems were of the same type, the chances were that the second problem was also correct. However, if the second problem was of the other type, the chances were that it was wrong. This suggests that Ss had a tendency to use exactly the same procedure as used on the previous problem, provided that the previous problem was answered correctly. A chi-square test was performed for each S and of the 64 Ss, 53 had significant chi-square values. Consequently, most appeared to have followed this procedure before they finally learned the concepts.

Test Scores

Defined as the total number of correct solutions given in the test booklet, test scores were highly correlated with both mental age and learning scores. The correlations with these variables were .72 and .85, respectively. A perfect score for this measure was 32. The mean score was 16.9 and the median was 15.5, with a standard deviation of 12, suggesting a large number of observations at either extreme of the distribution. Nineteen of the 64 Ss had a score greater than or equal to 30. At the other extreme, 12 of the 64 Ss had scores of 2 or less. The analysis of variance for this variable is presented in the first table. None of the F-ratios are significant except that for SES.

Note that while there was a significant difference in mean learning scores for the two grades, the corresponding difference for the test scores was not significant. The probable reason for this appears to be that the scores of those in the high SES 7th grade class reflected a ceiling effect; 12 of the 16 Ss had scores exceeding 26. If the test had been longer, they would have had higher scores. The mean test scores are presented in the following table.

Mean Scores on the Test as a Function of Grade and SES

SES	Grade		
	Fifth	Seventh	Both
Low	7.6	10.4	9.0
High	23.3	26.3	24.8
Both	15.5	18.3	16.9

The mean cumulative test scores for the four classes of students are shown in Figure 2 on the following page. As can be seen, the curves maintain their same relative position as they did for the learning curves. The only differences between the two sets of curves is that high SES 5th grade Ss perform almost as well as the high SES 7th grade Ss.

While the effect was not large, it should be noted that the experimental variables were not entirely without effect. For the learning scores the Variety by SES interaction was significant at $p < .10$. The source of this interaction centered on the low SES Ss. The mean performance for the low SES Ss was 16.4 for the small variety condition and 11.0 for the large variety condition. For high SES Ss the mean performances were 23.5 for the small variety condition and 22.5 for the large variety condition. For the test scores it is seen (in the first table) that both Variety and Form interacted with Grade and SES, through the four-factor interaction. According to the Scheffé Method of Multiple Contrasts, the significant source of variation is found within the 5th grade high SES Ss and the 7th grade low SES Ss. For the 5th grade high SES Ss, the mean score on the generalization test for the large variety condition was significantly higher for the nonsense forms than it was for the geometric forms, the average scores being 29.3 and 17.8 respectively. While the corresponding interaction for a 7th grade low Ss was significant it is not possible to make a clear statement concerning the effects of form upon variety for these subjects.

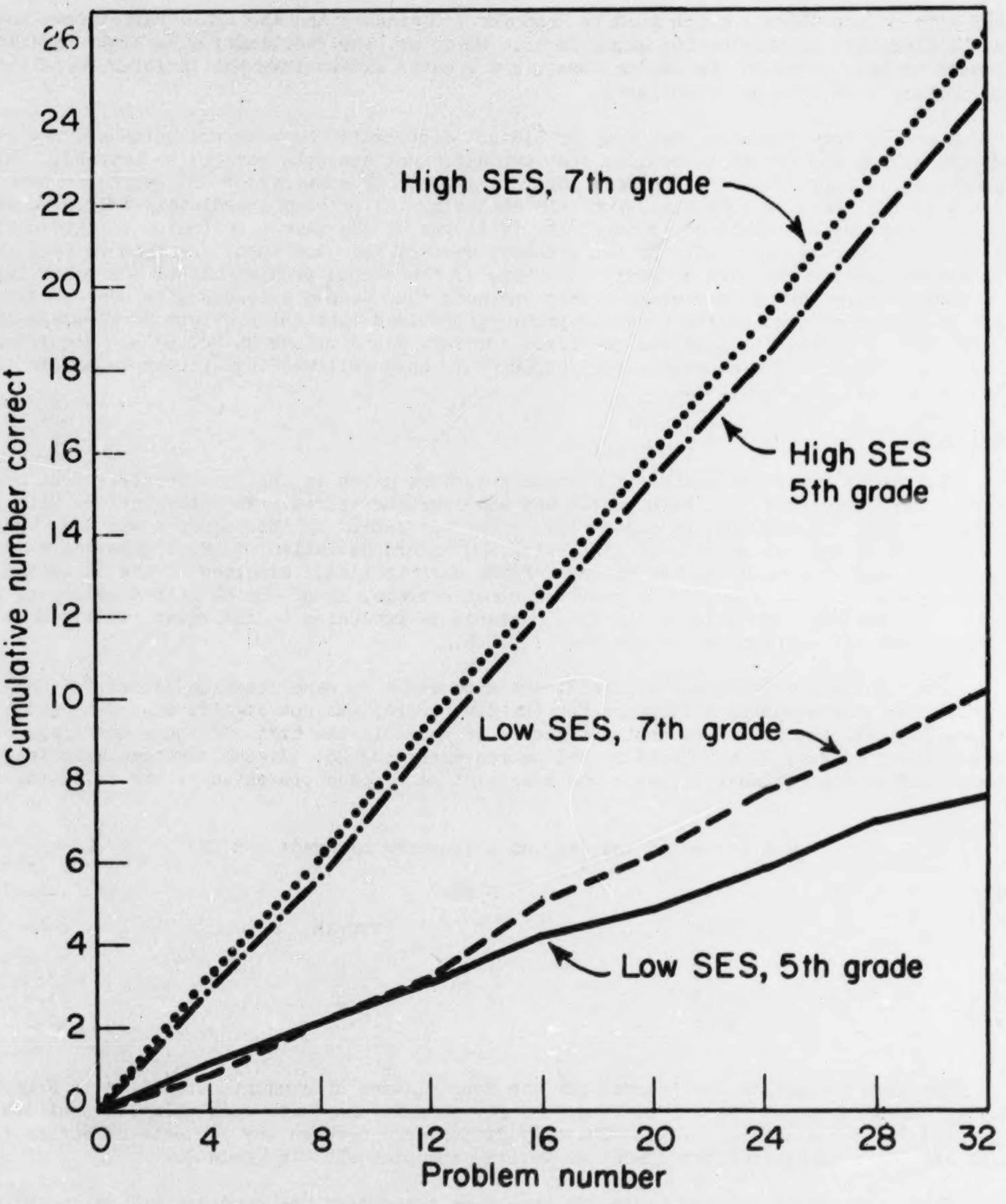


FIGURE 2. Cumulative Average Transfer Scores by Number of Problems. (Shown by blocks of four)

In sum, the results suggest that the variety variable has two significant effects. First, the small variety condition is more effective than the large variety condition for low SES Ss in learning the concepts in a generalizable way. Second, among 5th grade Ss the effect of variety depends on the type of form. For random forms the small variety tended to facilitate acquisition to a greater degree than the large variety while for nonsense forms, the means were in the opposite direction. The effect of variety is consistent with the result of the earlier experiment (Amster and Marascuilo, 1965) and suggests that it is limited to the lower SES individual. It is also consistent with a study by Amster (1963) employing verbal materials, in which the small variety was relatively beneficial for concept formation only among 4th graders who were slow learners in the experimental task.

The experiment described above was first performed as a pilot study with two different 5th grade classes in the same schools from which the two 5th grade classes used in the above analysis were located. For the pilot study 56 Ss were utilized in exactly the same experimental design. The only difference was that the instructions given to the Ss were briefer and not elaborated upon to as great a degree. Except for this, the experimental conditions were the same. The results were essentially the same. Most of the variability was related to the differences in SES status of the two schools or to differences among the children that attend the schools.

Summary

The present research was an attempt to extend the knowledge concerning the effects that variety of different experiences have upon the learning of a complex mathematical task. In paired-associate learning tasks it is known that any amount of variation is significant in aiding in the transfer to new stimuli. In learning the addition of integers, it seems to make no difference whether or not there is a low or high repetition of problems in transferring to new tasks. The results of this study support this conclusion.

Subjects were selected from 4 schools representing high and low SES schools. All Ss were in the 5th or 7th grade. For the experiment, Ss were divided into 4 groups. Every S was given 24 set union and 24 set intersection problems presented in a randomized block design. One group was given 48 different problems which used familiar geometric forms. The second group was given 12 different problems each repeated 4 times with the same familiar geometric forms. The third group was given 48 different problems that used unfamiliar nonsense forms. The fourth group was given 12 different problems each repeated 4 times with nonfamiliar nonsense forms. After learning this task, Ss were given 16 different set union and 16 different set intersection problems in which the stimuli were letters of the alphabet, simple English nouns, familiar everyday objects, and flags.

Analysis of variance on the learning and test scores showed that the most important variable associated with the learning task was the SES level of the schools. It is believed that a ceiling on the final test for the high SES Ss failed to produce an adequate test of the experimental variables. The differences between large and small variety experiences, nonsense and geometric forms, and their interactions were not significant at the .05 level. However, the effect of variety was noted to be consistent with an earlier study (Amster and Marascuilo, 1965) in which it was noted that the effect of variety was significant for lower SES Ss.

References

1. Amster, Harriett, The effect of the variety of instances on the production of verbal concepts. *Am. Psychologist*, 1963, 18, 382 (abstract).

2. Amster, Harriett, and Marascuilo, L. A., The effect of type of pretraining and variety of instances on children's concept learning. *J. exp. child Psychology*, 1965, Vol. 2, No. 2, 192-203.

3. Gagné, R. M., and Bassler, O. C., Study of retention of some topics of elementary nonmetric geometry. *J. educ. Psychology*, 1963, 54, 123-131.

4. Gagné, R. M., Mayor, J. R., Garstens, H. E., and Paradise, N.E., Factors in acquiring knowledge of a mathematical task. *Psychology Monogr.*, 1962, 76 (7, Whole No. 526).

Chapter 4

The Effect of Variety as a Function of Task Difficulty in the Acquisition of Word Meaning

The acquisition of word meanings from their use in verbal contexts is undoubtedly the most widespread manner by which verbal meanings are acquired and the series of studies reported in this paper concern possible ways in which children may do this. This problem is attacked by observing the efficiency with which verbal concepts are acquired in children of varying ages when the variety of different instances and the instructional set are manipulated experimentally.

In order to study the course of acquisition of word meanings in children, Werner and Kaplan (1951) devised an ingenious method, the Word Context Test, from which the procedures employed in the present studies were adapted. It involved presenting ten series of sentences, each series containing an artificial word. The same artificial word was used throughout one series and had a common meaning in all of those sentences. For example, the following sentences were among those from which the concept gather was to be abstracted:

All the children will lidber at Mary's party.

The police did not allow the people to lidber on the street.

The people lidbered about the speaker when he finished his talk.

The children were asked to define the concept as each succeeding sentence was presented. The first sentence remained in view when the second was shown. The authors attempted to analyze developmentally the reported changes in verbal meaning and found many changes with age, e.g., a striking drop in the occurrence of concepts which consist of parts of a sentence (sentence-contextual concepts) at about 10-1/2 years. They describe many types of concepts which occur and infer processes by which word meaning is derived from the verbal context. A common process is decontextualization, involving a gradual refinement in meaning in which the specific features of the contexts are dropped, until all that remains are the features of meaning which characterize the word in all contexts. However, this final stage of decontextualized meaning is only to be found in relatively mature individuals.

Bruner et al. (1956) analyzed the process of acquisition in the Word Context Test in a different way. They claim that the solution to each series of sentences is the word that has the highest probability across the series. This suggests that there is some associative process whereby the context probability of the new word increases from sentence to sentence. There is also a strong possibility that acquisition is a two-stage process involving not only association, but also deductive reasoning concerning whether the associations, formulated as hypotheses, "fit" the context. Thus, it seems reasonable that two processes of concept formation could be employed, as associative and a deductive process.

The extent to which hypothesis testing is employed in learning has been inferred from the extent to which learning is facilitated or inhibited by a relatively large variety of different instances (Osler and Trautman, 1961; Podell, 1958; Podell, 1964). Osler and Trautman (1961) have found an advantage of a small variety of instances over a large variety among relatively bright children. They reasoned that bright children generate more hypotheses from a large variety than from a small variety and therefore, the ability to select a simple hypothesis is hindered in that condition because of the interference from the many possible false hypotheses which must be rejected. On the other hand, the present author (Podell, 1958) has pointed out that the above consideration is only one of two factors which can influence the effect of variety. While the small variety should be relatively efficient with respect to the paucity of false hypotheses which they generate, the large variety should have an advantage over the small variety with respect to the efficiency with which any false hypothesis may be rejected (Podell, 1958). As described by Podell (1958), if a second example is the same as or highly similar to the preceding one, as it is likely to be in a small variety condition, the S probably cannot reject a hypothesis on the basis of the second instance which had been generated on the basis of the first. Since, for the large variety, a second instance is bound to be different from the first (or to be different in more respects than any two examples in a small variety condition), it is more probable that a S could reject in one trial any false hypothesis which he tested. Consequently, the extent to which a large variety should be more facilitating than a small variety should depend on the number of features within each instance, the number of potential associations which they elicit, the number of features common to different instances and the extent of their variation from instance to instance, and the absolute number of different instances in a set. Unfortunately, however, this type of standardization of instances remains to be done. In the present study, the definition of variety is the number of different instances in the set in which the instances are almost entirely different from one another but contain the same number of features. In this type of variety it is reasonable to suppose that the ability to reject false hypotheses is the most plausible reason for a difference between treatments differing in variety, and there is supporting evidence for this assumption (Podell, 1958). By contrast, when the complexity of the instances varies rather than the number of different instances, the volume of hypotheses to be rejected would by clearly greater in a large variety. Consequently, efficiency in relatively large varieties should be markedly lower than that in a small variety.

The use of deductive strategies in children's concept formation undoubtedly becomes increasingly prominent with age (Piaget, 1952; Inhelder and Piaget, 1958; Vinacke, 1954). The descriptive studies of Piaget (1952) and Inhelder and Piaget (1958), among many others conducted in their laboratory over the years, indicate an emerging dominance of mental operations of all types with age. The particular mental operations employed seem to change with age, culminating in the appearance of logical operations about age 11 to 14. On this basis, it would be expected that hypothesis testing as a type of logical operation would be employed in the acquisition of word meaning by children of about eleven years of age to a greater extent than it would be employed by younger ones. In addition, it is reasonable to suppose that these processes would be more prominent among relatively bright children, and evidence has been presented which has been interpreted as supporting the hypothesis (Osler and Trautman, 1961; Osler and Fivel, 1961).

It seems eminently reasonable that, if learned and available, logical reasoning would be more apt to occur under intentional concept learning than when no intent to form concepts or solve problems is present, and there is also evidence to support this assumption (Podell, 1958). Some consideration of each of the variables discussed is given in the experiments to follow.

Experiment 1 (Summary)[1]

A relatively easy Word Context Text was administered to 112 kindergarten Ss and 168 4th and 5th grade Ss, selected post-experimentally. Four concepts were acquired by children from their use in sentences. The two easy concepts were the nouns newspaper and train, and the two difficult concepts were the verbs gather and increase. A generalization test consisted of sentences containing the new word, which were to be accepted or rejected by the S. For intentional instructions, the scores on the large variety concept were significantly better than those on the small variety concept, $p < .05$ for the broad criterion and $p < .005$ for the narrow criterion. There was also a trend for the opposite result to obtain for the fourth and fifth grades given the Easy Concepts, and scored by the broad criterion ($p < .08$). The only significant generalization effects occurred in the fourth and fifth grade Ss. Although these Ss generalized more widely under the large variety, the interactions between this variable, variety, and Group limits the generality of the results. However, there was also a significant increase in generalization from the first to the second trial which was not similarly limited.

Sentence Standardization

The experiments which follow employ six sentences which contain one Cloze-type blank for each of ten concepts. Experiments 2, 3, and 4 which are to be described in succeeding sections of this chapter employed the same materials. The difficulty of these materials was assessed by the collection of norms based on isolated sentences, which consisted of the frequency of each of the responses elicited by the Cloze-type units.

---

[1] Supported under U.S.O.E. Project 1459.

Subjects

Sentence standardization procedures were administered in six 4th and 5th grade classes in the Berkeley Unified School District and the San Leandro Unified School District; data from 168 Ss was employed. The norms which are presented below are based on 28 responses to each sentence.

Materials

The materials consist of the six sentences employing each of the ten concepts as presented below. The single sentence norms are presented beside each sentence.

Standardization Procedure

Instructions were read to Ss in intact groups. A sample sentence was presented and discussed. The Ss were asked to listen to the sentence which contained a blank and to write down an English word which could go in it. They were told that each sentence would not be repeated and that they were to write down the first word that came to mind that seemed to fit in the blank.

From among the sixty sentences, each S received only twenty sentences from the set to be standardized, two corresponding to each concept. These twenty sentences were randomly interspersed among other sentences of a similar type in such a fashion that the entire procedure never required less than fifteen minutes. The six groups varied with respect to the particular selection of sentences and a different sequence of the entire set was employed in each group. The assignments were completely counterbalanced with 14 Ss in each subgroup.

Experiments 2 and 3[1]

Method

Thirty-two children from fourth and fifth grades were Ss for Experiment 2 and 52 Ss from the same grades were Ss for Experiment 3. The materials were as shown above. The same experimental design was employed for both experiments: A mixed design in which the repeated measure was Variety and the independently varied one was the assignment of concept to variety condition (group). There were two independent groups in each experiment, varying in the latter regard, and for Experiment 3 there were also two different random sequences in which the sentences were presented. For Experiment 2 the small variety consisted of two different instances each repeated three times, and for Experiment 3, the small variety consisted of three different instances each presented twice. For both experiments the large variety consisted of six different sentences. Sentences were read aloud to children tested in intact classes and they wrote their responses.

Results

Analyses of variance indicate that the effect of Variety tended to be significant ($p < .10$) for Experiment 3 but its generality is seriously limited by the significant interaction between Group and Variety ($p < .001$). The mean number of concepts attained under small and large variety were, respectively, 1.70 and 2.54 for Experiment 2, and 1.57 and 1.87 for Experiment 3.[2]

---

[1] Procedures are summarized since data collection for experiments 2 and 3 was supported under Project No. 1493. The analyses which entail single sentence standardization were carried out under the present project.

[2] The grouped sentences were re-scored for these comparisons and a broader scoring criterion was employed than in the earlier analyses. The mean scores were increased by less than 5%.

Comparison of Grouped Sentences with Single-Sentences

The relative frequencies with which each concept was acquired when the sentences were grouped in the 2-, 3-, and 6-sentence conditions were compared with three frequencies based on the frequency of emission of the concept derived from the single-sentence standardization. The median relative frequencies or estimates thereof based on each measure appear as follows:

Median Relative Frequency of Correct Responses in the Grouped Conditions
Compared with Values from the Responses to Single Sentences

Sentence Group	Grouped Obtained	Single (Expected)	T	Best Single Sentence (Expected)	T	Probability (Expected)	T
2 - sentence	.375	.161	4*	.232	10	.289	26
3 - sentence	.231	.208	17	.375	15	.524	5*
6 - sentence	.512	.197	1**	.500	12.5	.792	0**

* $p < .02$
** $p < .01$

The first column presents the median relative frequencies of the correct responses to the three types of sentence groups. This entailed obtaining the relative frequencies with which each concept was emitted by the grouped sentences for each concept under each of the grouped-sentence conditions. For example, for the concept *finish* 20 out of 26 *Ss* (.77) acquired the concept from the group of three sentences, and 13 out of 16 *Ss* (.81) emitted the concept from the group of two sentences, and 27 out of 42 *Ss* (.64) produced the concept from the group of six sentences. The median for all ten concepts is presented for each sentence. The second column contains the mean relative frequency with which each concept was emitted by the single sentences which correspond to those presented in the grouped condition. For example, the single sentence frequency for the concept *finish*, which corresponds to the grouped frequency for the two sentence condition, is the sum of the relative frequencies for each of the two single sentences employed in the grouped condition. Thus, 7 *Ss* out of 28 gave the concept in response to one of the sentences and 20 out of 28 gave the concept in response to the other. The relative frequency for the single sentence was 27 out of 56 (.48). Values were computed similarly for the 3- and 6-sentence conditions and for each of the ten concepts. The third column consists of the largest relative frequency of the response to the single sentences (Best Single Sentence). It was simply the frequency with which the concept was emitted in response to the one single sentence in the particular set for which the frequency was highest. In the above example, this frequency would be that of the sentence with the higher individual frequency, that for which 20 out of 28 *Ss* (.71) emitted the concept. The third value derived from the standardization frequencies was an estimated (Expected) value based on the assumption that the individual sentences have independent tendencies to elicit the concept and that an occurrence of the common response to any of the sentences is sufficient to evoke it in response to the set. The probability that the concept will *not* occur upon the simultaneous presentation of the two sentences would be:

$$\underline{P}_{comb} = (1 - \underline{P}_x) \quad (1 - \underline{P}_y)$$

In the example above, the probability that the concept *finish* would occur in response to the two sentences in combination would be:

$$\underline{P}_{comb} - 1 - \quad (1 - 7/28) \quad (1 - 20/28) = .79$$

Wilcoxin Matched-Pairs Signed-Ranks Tests were conducted between the grouped frequencies and each of the three corresponding standardization values. The assumption was made that the responses to the different concepts were independent. The corresponding frequencies for each concept were paired for these tests. The table presents the T values which correspond to each of the comparisons. More concepts are elicited by the grouped sentences than by the single sentences in all cases, but this fact is statistically significant only in the case of the 2- and 6-sentence conditions. There are no significant differences between the grouped-sentence conditions and the best single sentence and there is no consistent trend to the results. However, for the 3- and 6-sentence condition, the grouped frequencies are significantly lower than the expected values. Though not significant, the means are reversed for the 2-sentence condition.

Experiment 4

Subjects

Summer school children who had completed fourth grade and children who were beginning 5th grade were tested in intact classrooms in which all conditions were run simultaneously and in which Ss were assigned to conditions by random means. From among those tested, 108 Ss, 36 from each condition, were postexperimentally selected with a view to excluding Ss known to be repeating fourth grade. Random means were employed in postexperimental selection. San Leandro and San Lorenzo Unified School Districts provided the Ss for this experiment.

Materials

The ten concepts, each exemplified by six sentences which were used in the two preceding studies, were employed.

Experimental Design

The major variable under investigation was Variety, the number of different instances from which the concept was to be identified. The three variety conditions were small variety consisting of 3 different exemplars of each concept, small variety repeated, consisting of 3 different exemplars, each repeated, consisting of 6 different exemplars. Independent groups were employed. One of the control variables, arrangements, was the particular sequence in which the ten concepts were presented, and there were three different random sequences, equally represented by 12 Ss within each variety condition. Three assignments of nonsense name to concept were employed and deliberately confounded with arrangement. The other control variable was the particular set of six sentences from within the large variety which was chosen for the small varieties and the sequence in which the set of sentences for each of the concepts was presented, called Subarrangements. The design was completely counterbalanced over each small variety such that each sentence appeared an equal number of times within the set of instances administered under each variety condition; thus, equating the frequency of occurrence of particular sentences across variety conditions. Two Latin Squares were used to assign sentences to Ss within conditions. The sequence in which the sentences appeared was also counterbalanced in such a way that each sentence appeared an equal number of times in each possible position within each variety condition. Booklets were matched over variety conditions in the following manner: if a particular set of three sentences was chosen to exemplify a small variety concept and presented in the order 1, 2, 3, they were presented in the corresponding small variety repeated condition as 1, 2, 3, 1, 2, 3, and in the large variety condition as 1, 2, 3, 4, 5, 6. Similarly, 4, 5, and 6 appeared in the first case in the given order, in the second case, as 4, 5, 6, 4, 5, 6, and in the final case, as 4, 5, 6, 1, 2, 3.

Instructions

Don't turn your booklets over until I tell you to do so. When you do, you will find that each page has some sentences with a new word in them. All the sentences on one page have the same new word like the sentences on the board. (Write these sentences on the board in advance:

(1) Most children have a father and an INDAR.
(2) The baby loved his INDAR.
(3) INDARS usually make the meals for the family.)

There will be a different new word on each page. Your job is to read the sentences one at a time starting with the first one. (Read) then write down your best guess as to what the new word means on the line at the bottom of the page. If the line is not there, just write your answer on the bottom of the page somewhere. What do you think INDAR means? (Write the answer on the board as they should write in their booklets.) If you are not sure, but have an idea of what the word could be, write it down. If you really don't know, leave it blank and go to the next page. *Don't read the list of sentences over again!!* Suppose you read the sentences on the board and didn't know what the word INDAR meant after you read them once, what would you do? (Let them answer and say RIGHT when someone says, "Turn the page", and WRONG to other answers. Repeat what they should and should not do.) Remember, read all of the sentences on a page only once. If the same sentences appear again, you should read them each time. Suppose there were four sentences on the page. (Add the baby loved his INDAR as a fourth sentence.) You would read the four sentences in a row (read), write the answer (write) and then turn the page.

Any questions? When you are finished, turn your booklet over like it is now. Work quickly, but read carefully. Before you begin, write your name and age and school on the back of the booklet.

Results

The data were scored according to two criteria, a *strict criterion* by which only those responses which were clearly correct as a response to all six sentences comprising a large variety were considered correct, and a *lenient criterion* by which responses were judged in accordance with their correctness or plausibility in response to the set of sentences which that particular *S* received. For the large variety, the same scores obtained for both methods of scoring, but for the small varieties, more concepts were identified correctly when the lenient criterion was employed. The table presents the mean number of correct concepts for each variety condition and each scoring criterion. For both scores, the means corresponding to the two small varieties are almost identical and considerably lower than the mean for large variety condition. However, analyses of variance reveal a significant effect of variety for the strict criterion ($F = 5.25$; df = 2/72; $p < .01$) while the corresponding effect for the lenient criterion did not reach an acceptable level of significance ($F = 1.72$; df = 2/72). However, a further analysis of the lenient scores was undertaken, which did indicate that variety had a reliable effect. For this analysis, separate scores were obtained for each *S* for the easy and difficult concepts. The following table indicates the total number of correct responses to each concept given separately for each variety condition. Concepts were classified as easy and

Frequency of Correct Responses (%) According to Strict and Lenient Criteria and Expected Frequencies (%) Based on the "Best Single" and Mean Dominance for the Strict Definition of the Expected Values[1]

	Best Single	Mean Dominance	Strict	Lenient
Small Variety	50.03	28.81	38.3	46.7
Small Variety Repeated	50.03	28.81	36.9	47.2
Large Variety	65.00	28.81	52.5	52.5

Mean Number of Correct Identifications of Easy and Difficult Concepts for each Condition of Variety[2]: Lenient Scoring and Lenient Definition of Difficulty

	Easy (n=5)	Difficult (n=5)
Small Variety	2.89	1.78
Small Variety Repeated	2.94	1.80
Large Variety	3.50	1.75

---

[1] In deriving expected values for the "strict" definition, the "strict" frequency was used.

[2] Difficulty was defined separately for each variety condition. See appendix for details.

difficult separately for each variety condition on this basis. The analysis of variance again indicates the main effect of Variety to be nonsignificant ($F = 1.91$, d.f. 2/72, $p < .05$), but the interaction of Variety and Difficulty was statistically significant ($F = 3.68$, d.f. 2/72, $p < .05$). The effect of difficulty was of course, highly reliable ($F = 107.06$, d.f. 1/72, $p < .001$) and the effect of subarrangements, as in the other analyses, was also highly reliable, ($F = 4.68$, d.f. 5/72, $p < .005$). The latter indicates that the particular set of sentences presented to exemplify the concept and the sequence in which they were presented significantly affected difficulty.

Mean Number Correct (%) of Easy and Difficult Concepts for
Which (Strict) and Lenient Scores are Presented for
Strict and Lenient Definition of Easy and Difficult[1]

	Lenient Definition		Strict Definition	
	Easy	Difficult	Easy	Difficult
Small Variety				
Lenient	48.4	28.4	57.6	35.6
Strict	(57.6)	(35.6)	(48.4)	(27.8)
Small Variety Repeated				
Lenient	49.4	24.4	57.2	37.8
Strict	(58.8)	(36.2)	(48.8)	(24.4)
Large Variety				
Lenient	70.0	35.0	70.0	35.0
Strict	(70.0)	(35.0)	(70.0)	(35.0)

[1] Lenient definition of easy entailed the five easier concepts and five harder selected separately for each condition. Strict definition consisted of easy and difficult being selected according to the large variety and used for all conditions.

The first table in this section includes the expected frequencies of correct responses, based on the expected strict responses which should occur in response to single sentences. These are presented in the appendix. The definition of easy and difficult used in the second table and in the analyses of variance is the one in which a separate determination of easy and difficult concepts was made for each variety condition by observing the number of correct responses to each concept under each condition. However, the means relative to a strict definition are also presented in the third table. For this definition, the same classification of concepts into easy and difficult which obtained for the large variety condition was used for all conditions. The differences between variety conditions with respect to their expected difficulty may be compared with the obtained frequencies.

The strict scores (%) were analyzed further to determine the extent to which the obtained scores deviated from "Best Single" expected scores. These expected scores were the frequency in percent with which the standardization Ss "correct" concept elicited in response to the single sentence in the set to which it was most dominant. Thus, the expected means were the same for all Ss receiving a large variety, but they varied from S to S under the small variety conditions since the particular best single sentence varied as a function of the specific subset of sentences administered. The most striking result was that for all conditions, the obtained values were reliably below the expected values. In each of the small variety conditions only one out of the respective 36 Ss obtained scores which were above the "best single" scores and in the large variety condition 7 out of the 36 Ss obtained scores above the best single and these frequencies were reliable on the basis of Sign Tests ($p < .01$). An analysis of variance on the difference score for each S (percent correct using the strict criterion minus the percent correct estimated from the best single response) was conducted in order to determine differences as a function of conditions. However, although the large variety conditions showed a smaller deviation from the expected values than the small variety conditions, the main effect of variety was not significant ($F = 1.82$, d.f. 2/72, $p < .25$). As in prior analyses the main effects of Arrangement and Subarrangement were significant ($F = 13.39$, d.f. 1/72, $p < .005$ and $F = 2.76$, d.f. 5/72, $p < .05$). For this analysis, Arrangement referred to the two Latin Squares which were used to counterbalance the particular selection of sentences and their sequence of presentation. The interactions of these control variables and variety were not significant. The mean differences (%) for the three variety conditions may be assessed from the first table in this section.

Discussion

Taken together, the experiments indicate that for practical purposes many different sentential contexts are more likely to elicit the correct concept than fewer sentences drawn from the same set. However, progressive refinement of the analyses undertaken and the controls employed revealed certain limiting conditions under which this was true. Specifically, the large variety was significantly superior to the small variety when performance was measured in terms of a strigent rather than a lenient criterion i.e. when the same criteria for correctness was employed for small and large varieties. It also tended to be superior otherwise. Moreover, the large variety was reliably superior even for the lenient criterion when easy concepts were considered, but there was hardly a detectable difference between variety conditions for the difficult concepts. One might think that there was some summation upon increasing the number of dominant contexts, but none upon increasing the number of remote contexts. However, the group of sentences was no more likely to elicit the concept than the best single sentence with set, i.e., analyses of the relationship between the obtained frequencies of correct responses and the standardization frequencies in response to the best single in the particular set revealed for experiment 3 that there was no reliable difference between obtained and expected values although Ss tended to perform below expected values, and for Experiment 4 that obtained stringent scores were reliably below the best single under all conditions. Although the values tended to be closer to the best single under the large variety, the effect of variety was not significant. For Experiment 2, the small variety of two instances did tend to stimulate the correct response to a degree in excess of the best single, but this may have been an artificat. The effect of a small variety of three was replicated in two experiments in which the occurrence of correct responding was below the best single. Thus, clearly, Ss are not merely summating; their scores in response to a group of instances do not even equal the frequencies which would occur in response to the best instance in the set. But perhaps, it might be wisest to conclude that any summation which occurs is usually offset by the interference which occurs from successive instances.

The fact that obtained values were significantly better than mean dominance was established for both experiments, but no summation was observed relative to the "best single" response. Undoubtedly, Ss received higher scores in Experiment 4 than in the previous experiments because auditory presentation was used for the former.

It would appear that the ease with which a set of instances will elicit a concept depends mainly on the probability with which the most dominant instance in the set will elicit that concept and not on the sheer number of instances in the set. This was demonstrated in part by the fact that the large variety was only facilitative for easy concepts and not for difficult ones. The sheer number of instances has an effect in that it provides a greater opportunity for the occurrence of highly dominant instances than does a small variety, on the average. But if their probability of occurrence is low, the amount of interference must out weigh the possible summation. However, it is likely that hand-picked small varieties which include the most dominant instances which are included in any large variety set would produce better results than the large variety which induces more interfering responses than the small variety. It is quite clear that in all experiments the biggest factor affecting difficulty was the particualr set of sentences selected, and inextricably confounded with it in the present experiment, the sequence in which they appeared. The particular assignment of nonsense words to concepts and the order in which concepts were presented appeared to have very little effect.

The final experiment compared a small variety of three with a small variety of three repeated, on the hypothesis that if repetition of the same instances generated new hypotheses with the same probabilities that new instances do, the small variety repeated would function as a large variety. However, there was barely any detectable difference between the two small variety conditions in any analysis. In other words, the effects were almost perfectly replicated in the two variety conditions. They suggest that concept formation does not depend merely on the number of occasions for generating hypotheses, but rather, on the extent to which the succeeding instances have high probabilities of eliciting the correct hypotheses and have not been observed previously. It is as if each S can only make one associative response to each instance and thus it is harder to make a new associative response to the same instance than to make (new) responses to a new instance.

Chapter 5

The Effect of Variance and Variety in Convergent Association

Prior Work on Convergent Associations[1]

As part of a prior study, (Amster and Keppel, 1966), norms were collected for stimuli which consisted of pairs of words which had been used as stimuli on the Palermo-Jenkins list and were known to have at least one response in common. The responses which were collected for convergent association norms consisted of the first word that came to mind after reading both words of the pair. As a way of predicting the frequency of convergent responses, the mean response dominance of one common association was computed as the mean frequency in percent with which it occurred in the single word norms, and the convergent frequency of that response was assessed from the convergent norms which were obtained. On the basis of considerations concerning two processes which might contribute to the production of convergent associations, associative summation, and hypothesis testing, it was expected that the frequency of convergent associates would depend not only on the mean dominance of the common convergent response, but also on the variance between the two members of the pair with respect to the frequency with which each member tends to elicit the predicted response.

The convergent association task had been selected as a simplified miniature concept formation situation and the variable of variance among stimulus members is one which is ordinarily allowed to vary at random among sets of instances provided for concept elicitation. Specifically, but depending on how they were selected, a small variety of instances would have a smaller variance than a large variety of instances. The convergent association situation permits variance to be studied while holding constant the number of different instances within a set.

As shown in Table 1 the results of the convergent association concerning mean dominance were as expected, i.e., for both adults and children, the frequency of convergent associations increased directly with mean dominance. However, the results for variance were generally small and insignificant although for high dominance pairs, children produced the selected convergent associate significantly more frequently for low than for high dominant items. The same trend, though nonsignificant and smaller, was found for adults. This finding could have come about through the checking of hypotheses if children rejected associates on the basis of inability to find some relevance between them and the less dominant member of the pair. Perhaps they were less efficient at seeing the relatively remote relevance of any association produced than were adults in a comparable situation. On the other hand, it might be that greater summation occurred for children in response to low variance pairs than occurred for adults. In support of this hypothesis, it was found, as shown in Table 1, that for low variance pairs, the amount by which the frequency of the convergent associate exceeded the expected values based on mean dominance was considerably greater for children than for adults, while for the high variance pairs, the adults' performance differed little from the expected value while the children's was clearly below it. Consequently, support was found for postulating that both processes contributed to the obtained convergent strength in ten-year-old children.

The Effect of Variety in Convergent Associations

A major factor believed to affect concept formation and the production of convergent associations is the amount of interference from responses competing with the desired response. In the usual concept formation situation where variety is studied, the mean dominance of the instances and their variance may be allowed to vary at random, thus

---

[1] Collection and processing of these data were supported by a grant from NICHD while collection and processing of the data of the next section in this chapter were supported by the present project.

obscuring the study of competition. The convergent association situation however, permits ready control of both of these factors. The preceding section sketches briefly a study which provides evidence that these variables do affect the production of convergent associations and suggests how they enter into an uncontrolled selection of a large and small variety of instances. In the present section, an attempt is made to consider interference and summation in a situation where differential results due to hypothesis checking are minimized by equating for variance, sets of stimuli which vary in the number of items which compose them.

If one again considers the convergent association situation as a miniature concept formation situation, the effect of variety can be analyzed by manipulating the number of different stimuli in the set from which the frequency of the convergent response (s) to the set are to be measured while controlling for variance. As was evident from the studies of the acquisition of verbal concepts, large and small varieties differed with respect to the strength of the most dominant instance in the set; the most dominant instance invariably being present in a large variety and rarely present in a small variety. Similarly, the preceding section indicates that not only is the mean dominance of the set of stimuli an important variable, but also the range of dominance or variance of the set. It should become apparent that this range would be greater on the average in a large variety than in a small variety of equal mean since the most deviant instances would invariably be present in the large variety which contained all (or most) of the instances in the set. Consequently, it seemed desirable to determine the effect of variety in a situation in which the stimulus sets were controlled both for mean dominance and for mean variance.

Assuming that mean dominance and mean variance were controlled, it would be expected that children would be more likely to produce convergent associates in response to many different instances than a small number of different instances since each additional instance having a tendency to elicit the response would add to the strength of the cumulative tendency. On the other hand, each additional instance would multiply the number of competing associates since all the other associates to that item might be elicited and all the other associates common to that item and the others in the set might be facilitated. Furthermore, a correct association might be rejected on the basis of successive instances if the S is testing hypotheses and can find no relevance of the association to the new instances. However, these two considerations could be considerably reduced in children compared with adults who presemably have more associations to each item and thus have more potential competitors. Also, adults may have stronger tendencies to test hypotheses.

In an earlier study, (Podell, 1963) conducted with college age adults, fewer convergent responses tended to occur to quartets (sets of four words having a common associate) than to pairs (sets of two words having a common convergent associate). The most likely hypothesis was that adults suffered considerable interference from competing associations which increased exponentially as the number of instances was increased. Quite possibly children would suffer less from such competition.

Subjects

One hundred of the fourth grade Ss to whom convergent association pairs had been presented were given the present materials, following their responses to the pairs. They were from the San Lorenzo Unified School District in California.

Materials

Thirty two sets of three words (triplets) and thirty sets of four words (quartets) were selected of which all were stimuli on the Palermo-Jenkins list and for which there existed at least one response in common for each word comprising a set. The fourth grade responses from the Palermo-Jenkins norms were employed for this selection. Further, the triplets and quartets were selected from among equivalent ranges of dominance, and stimulus sets which could be construed as homogeneous in part of speech. Triplets and quartets were roughly matched in these respects.

The triplets and quartets were interspersed and appeared in one random order in half the booklets and in its reverse in the other half of the booklets. The order of presenting the words within the stimulus set was randomly determined and differed for the forward and backward versions of the stimulus lists.

A complete list of the stimuli from which the items were selected for the comparisons which follow appears in the Appendix. It includes the mean *dominance levels* (mean frequency in percent on the Palermo-Jenkins norms of the convergent response to each stimulus, the variance defined as the difference in dominance between the most dominant and least dominant stimulus in the set, and other pertinent information.

Procedure

Ss numbering 104 from three classes were assembled in a large auditorium. The booklets were randomly distributed to *Ss*. Instructions were read aloud. Words which the children could not read were read to them by the experimenters or the teachers who were present. There was no restriction on the time to complete the task, but they were urged to work continuously. Two *Ss* refused to cooperate; two others were discarded at random from the other conditions.

Instructions

"Don't turn your booklets over until I tell you to do so. I am going to try to tell you what to do before you actually look at what is in them. When you do look, you will see that your booklet has pairs of words on the first page. No don't turn the page until you have finished working on the first page and then finish the second page before you start on the third page. You will be allowed to turn the page yourself as soon as you finish any page. You will find that the first and second pages have pairs or sets of two words on them. The third and fourth pages have sets of three words and sets of four words. Now this is what you do: READ *BOTH* WORDS AND WRITE DOWN THE FIRST WORD THAT COMES TO MIND AFTER YOU HAVE READ *BOTH* THE WORDS. For instance; consider this pair of words;

Cat
Lady ______________

You would read both words (cat, lady) and write the word that comes to mind. What do you think of? (Get them to respond and say OK after each response). Notice that you can write any word you think of, if it is the first word you think of after reading both the words. You might write *fur* (write) or *friend* or any you think of. Notice that the word you write should be a response to both *cat* and *lady*. THERE IS NO RIGHT OR WRONG WORD! Suppose there were three words in the set like this:

Horse
Tree
Foot ______________

You would do the same thing you do with two words. Read all three words and then write the first word that comes to your mind. Raise your hand if you don't know what to do. Work quickly, and try to finish before the hour is over, but be careful to read all the words in each set before you think of another word, and do not skip any blanks. Try not to look at your neighbor's booklets. Also, try to write or print very clearly -- we don't care how you spell the word, but we want to be able to read every letter."

Results

The percentage frequency of occurrence of the specified convergent associate in response to each set of stimuli is shown in Table 2. It may be apparent from this table that many different sets of triplets and quartets may be selected which match within narrow limits on certain variables but not others. Several sets matching in mean dominance, variance, and part of speech were selected. In addition, a set of pairs which match the triplets and quartets in mean dominance and mean variance was also obtained. The dominance levels, variance means, and part of speech of each set are presented in Table 2, together with the frequency with which the convergent association occurred in the convergent association test, in each set of stimuli. Significance of the differences were assessed by means of a sign test for which each *S* was scored for the number of convergent responses he produced in response to each set. The adjective triplets and quartets did not differ

TABLE 1

Mean Convergent Frequencies (%) for Adults and Children, of Responses Varying in Response Dominance (%)[1]

Dominance	Extremely Low		Very Low		Low		Medium		High		Very High	
					ADULTS							
Range of Response Dominance %	0.4		2.5		2.6 - 5.2		5.2 - 10.0		10.0 - 20.0		20.0-30.0	
Variance			Low	High	Low	High	Low	High	Low	High	Mixed	
Convergent Frequency %			2.31	2.98	11.57	10.05	12.70	15.61	15.74	14.15	22.09	
					CHILDREN							
Response Dominance %	0.4 - 1.2		1.4 - 2.4		2.6 - 5.0		5.0 - 10.0		10.0 - 20.0		20.0-30.0	
Variance	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
Convergent Frequency %	1.29	1.39	4.96	6.15	7.74	4.17	9.08	9.67	20.19	10.62	24.06	23.56

---

[1] For adults 6 pairs represented each cell; 54 pairs in all. For children, 8 pairs represented each cell; 96 pairs in all. Response dominance based on Palermo-Jenkins norms convergent frequencies based on 252 adults and 252 children.

in the frequency with which they elicited the convergent associate. The medium and high dominant noun quartets elicited the convergent associate in significantly more Ss than did the triplets. An analogous result was obtained for sets composed of mixed parts of speech, but it must be kept in mind that the latter were largely composed of the same stimuli. For the sets composed of mixed parts of speech, pairs of approximately matching mean dominance and variance were also obtained. For low dominant sets the pairs elicited the convergent response in significantly more Ss than either the triplets or quartets. For medium dominant sets the pairs elicited the convergent response from significantly more Ss than the triplets but they did not differ from the quartets.

TABLE 2

Sets of Pairs, Triplets, and Quartets Which are Matched for Mean Dominance and Mean Variance and Specified with Respect to Part of Speech

	PAIRS			TRIPLETS			QUARTETS		
ADJECTIVES	Mean Dom.%	Mean Var.%	Convergent Freq.%	Mean Dom.%	Mean Var.%	Convergent Freq.%	Mean Dom.%	Mean Var.%	Convergent Freq.%
(n = 7)				11.57	20.31	12.71	13.01	19.66	12.00*
(n = 6)				11.31	20.63	11.83	10.50	18.83	8.83*
NOUNS									
Low Dom. (n = 7)				2.97	3.03	2.57	2.88	3.06	2.57*
Med. Dom. (n = 7)				6.81	11.63	6.29	6.96	11.74	10.00[1]
High Dom. (n = 7)				13.98	23.17	8.43	13.96	29.57	14.00[2]
MIXED PARTS OF SPEECH									
Low Dom. (n = 8)	3.05	2.65	4.63	3.14	3.32	2.63	2.85	2.80	2.50[3]
Med. Dom. (n = 8)	6.75	9.60	9.38	7.03	12.00	6.28	6.99	11.18	9.38[4]
High Dom. (n = 8)	14.86	22.52	11.75	13.72	21.65	9.50	14.25	30.28	13.75[5]
Overall (n = 24)			8.58			6.17			8.54

* Triplets and quartets were not significantly different in convergent frequency.

[1] The frequency for quartets was significantly higher than for triplets ($p < .02$).

[2] The frequency for quartets was significantly higher than for triplets ($p < .01$).

[3] The frequency for pairs was significantly greater than for triplets ($p < .05$) or quartets ($p < .02$); triplets were not significantly different from quartets.

[4] The frequency for pairs was greater than for triplets ($p < .05$) but not significantly different from quartets; quartet frequency was significantly higher than triplet frequency ($p < .01$).

[5] The frequency for pairs was not reliably different from the others; the frequency for quartets was significantly higher than for triplets ($p < .01$).

Since the results for the matched sets of triplets, quartets, and pairs were inconsistent, the data were reanalyzed. All available data were plotted (not shown). Instead of using the convergent response for pairs which was previously used, the most frequent common response to each pair was determined in order to match them more closely to larger sets, i.e. since common responses to three or more items are so rare, the common response which had been employed was thus the most frequent. Thus the pairs equated with triplets and quartets were plotted and showed a clear superiority of pairs over triplets and quartets which seems to increase markedly with frequency. The figures (not included) show the frequency of convergent responses as a function of the mean dominance of the set and as a function of the most dominant word in the set (best single). Curiously enough, both figures, highly similar, reveal the triplets to be less facilitative than the quartets, and this curvilinear relationship recurs consistently in the comparisons which have been made, but the values for the two largest sets are extremely close.

Since the figures discussed above do not involve controls for variance in the case of the mean dominance curve or mean dominance in the case of the curve of the most dominant word, another plot was made in which sets of pairs, triplets, and quartets were matched item for item very closely for the dominance of the most frequent common convergent response and within availalbe limits, for the frequency of the best single response. Triplets and quartets were matched first. Their dominance fell within 10% and when the frequency of the "best single" was unequal, it was greater for that item having the lower dominance. After 14 matched pairs of triplets and quartets were assembled, pairs which matched the weaker of the two as closely as possible on both measures were collected. The data are presented graphically in the appendix.

Although the convergent frequencies did not differ among the three sets on the basis of a sign test, the mean convergent associations obtained differed considerably in freqency (%). The mean for pairs was 14.20, for triplets, 6.36, and for quartets, 9.64; the superiority of pairs over the larger sets was marked. Superiority of pairs was significant ($\underline{F} = 3.50$; d. f. 2/26; $\underline{p} < .05$).

Discussion

In an earlier study of adults (Podell, 1963), the frequency of convergent responses to pairs were found to be greater than the frequency of convergent responses to quartets which were matched to the pairs in combined probability of eliciting the predicted response and were also superior for pairs equal to the quartets in mean dominance when the variance was low. For the present study of ten year old children, a similar superiority of pairs over quartets (and also over triplets) was observed. When the convergent response which was employed was the most frequent common response on the basis of the single-word norms. However, when the response was unselected, as for adults, no consistent superiority of pairs over the triplets and quartets was observed. This trend toward less difference between pairs and quartets for children is consistent with the notion that fewer competing responses exist for children than for adults. A striking difference among the three sets was that the pairs were clearly the most variable.

Wherever a difference was observed, the quartets were consistently superior to the triplets. However, the magnitude of the difference was considerably smaller than the difference between either of those sets and the pairs. The clear superiority of the pairs to the other items is quite consistent with the results for adults and with an interference interpretation of our data in that it would be expected that the number of responses which could compete with the common convergent response might be much greater in sets of three or four than in sets of two because the number of associates which could compete might be estimated on the basis of the total number of available associates to all of the single words in the set. In addition, more different responses are given to compounds than to single words (Musgrave, 1958). Thus, the larger the set, the larger the possible pool of words which could compete with the predicted response. Clearly, the possible summative effect of adding stimuli is minimal compared with the detriment which derives from adding to the number of potential interfering responses. Apparently, however, the balance shifts as the number of items added to the set increases. The small but consistent advantage of quartets over trios suggests that the number of new competitors introduced by the addition of one new stimulus when three are already present is completely offset by the gain due to the additional elicitor of the response. Consequently, there is reason to believe that both factors are operative.

A further reason for believing interference rather than summation to the major determinant of obtained differences is the fact that the observed convergent responses for pairs were of about the same strength as that predicted from the "best single" stimulus, but did not exceed this value. On the other hand, for triplets and quartets, the convergent frequencies approximated the mean dominance of the set and were markedly lower than those expected on the basis of the "best single" response in the set. In other words, there is no evidence of any induced elevation in the frequency of the convergent response for sets of any size. But interference was greater in the larger sets.

Another factor affecting the superiority of pairs concerns the relative influence of any one item in a set. If the set is small, it would undoubtedly bear a large proportion of the total weight, compared to its effect in a large set. But in addition, there is reason to believe that the first and last items in any set have greater influence than middle items in determining the total value of the set (Amster, 1966) and needless to say, the probability that an item will appear in one of these positions is greater when there are only two items in the set than when there are more than two.

Chapter 6

The Effect of Convergent Associative Strength and Stimulus Variance on Conceptual Grouping and Learning Difficulty

Frequency of convergent responses should depend not only on dominance, but also on variance. If dominance level and variance determine convergent frequency, they should also affect learning rate in a similar fashion. Consequently, pairs which were known in those respects were selected for paired-associate learning. It was thus expected that for high and low levels of dominance, low variance pairs should be learned more readily than pairs of high variance but rate of learning should also increase directly with level of dominance. Strength of the associative connection between the stimuli and the responses should be a major source of pair difficulty. In contrast to children, for adults strength of associative connection between words has been found to have little effect within a broad range. Carlin (1958) found that all degrees of association produced significant mediational facilitation, but they did not differ in relative efficiency. However, degree of associative strength has been shown to affect difficulty of learning word-pairs in children of three grade levels (Shapiro, 1965). The present study attempts a further investigation of this factor.

The nature of convergent association suggests the relevance of learning of pairs within doublets such that a convergent response would be learned to each of the stimuli which elicited it. This would make for the same response being learned to two stimuli. If the *S*s are learning such lists by doublets instead of by single pairs, it would be expected that in addition to response dominance, convergent frequencies should predict learning difficulty and the tendency to group pairs into doublets during learning. This tendency to group words together on the basis of some conceptual similarity should also be a function of the strength of any strong common convergent associate to them, irrespective of whether that common convergent associate is the one being trained. For example, assume that the convergent primary to *black* and *white* was *color* and this was a strong (highly frequent) response. Further, assume that the response to-be-trained to the two was *dark*; *black* and *white* might tend to be grouped together and the response learned more readily to *black* and *white* as a unit than would occur in another doublet in which the stimuli did not have some common associate which would tend to link them. On the other hand, it might be more difficult to learn a particular convergent response if the strength of the connection of the convergent response to the individual words was weak and the strength of the convergent primary was relatively strong and thus competed with the learning of the other convergent response. Consequently, the tendency to group pairs within doublets during learning should be assessed as a variable separate from speed of learning, and strength of convergent primary could influence them differently.

Strength of primary might be expected to influence speed of learning in a way which is diametrically opposed to the hypothesis offered above. It has been found for the learning of single pairs of words, that ease of learning depends directly on the strength of the associative primary responses. To the stimulus words in that responses are more easily learned to stimuli which elicit strong primary responses in free-association tests than to stimuli which elicit relatively weak primary free-associates. This phenomenon was first noticed in the control conditions of an experiment reported by Palermo and Jenkins (1964). In that experiment, ease of learning the control pairs varied as a direct function of the strength of the primary normative responses to the stimulus words. The effect occurred again in the control conditions of an experiment conducted by Wicklund (1964), although a reversal in the expected trend was found for the fourth grade *S*s. A supplementary study (unpublished) was conducted at that time using *S*s from the same classrooms as those in the discrepant control group and employed a nine-pair list of unassociated words. The stimuli of three of the pairs elicited very strong normative primaries, three at an intermediate level, and three elicited relatively weak primary responses. None of the stimuli in this list were the same as those for the list presented to the discrepant group. The findings for the supplementary group supported the original finding, i.e., strength of primary response, even though that response does not actually occur in the experimental situation, varied directly with ease of paired-associate learning.

A final replication (unpublished) was performed to rule out the possibilities that the earlier lists were confounded by some orderly variation of any traditionally effective variable. Stimuli eliciting either strong or weak primary responses were paired with unassociated response words. All words were nouns of four or five letters with Thorndike-Lorge classifications of A or AA. Four different sets of re-pairings of stimuli and responses were used; and, since the phenomenon had always occurred under mixed-list conditions, mixed-list and unmixed-list conditions were also compared. The findings again replicated the earlier outcomes and supported the contention that the phenomenon is not artifactual.

The present study is an attempt to determine whether a similar effect is found when the strength of primary variation is based on Convergent Association normative materials. Since the Convergent Association norms were collected in a free-association situation, the associative mechanisms should be highly similar and resulting associative materials should have characteristics similar to those of regular, single-stimulus, single-response norms.

It should be noted that in the present experiment the response words are associated with the stimuli whereas this was not the case in the studies discussed above. A possible alternative explanation for the effect has been suggested by Wicklund, Palermo and Jenkins (1964) for situations in which associates are learned. The outcome of the present experimentation might suggest some resolution of the alternative explanations noted above.

It is possible that the strength-of-primary phenomenon is a reflection of competition between the response to-be-trained and the primary. A strong primary may be more readily discriminated from a response to-be-trained than a weak primary and this factor could operate similarly with respect to the learning of doublets. Thus, a strong convergent primary should facilitate the learning of doublets for which the response was a convergent associate compared with a weak convergent primary.

The experiment described in the preceding chapter indicated that for ten-year-old children, the production of convergent associates to pairs was more frequent than the production of these associates to triplets or quartets when the overall dominance level of the convergent associate was low. But when the dominance level was increased, quartets were more effective. This suggested that the relative importance of two factors, amount of interference-to-be-overcome and the absolute levels of dominance of the constituent stimuli, might shift in importance. The amount of interference would be expected to shift with level of dominance, in that with increasing mean dominance, the convergent responses being studied would be higher in the convergent hierarchies. Otherwise, on the average, the same number of competitors would obtain, for pairs of high and low dominance, but for all levels of dominance, the number of potentially interfering responses would increase with the number of stimuli in the set. Thus, it appeared that for low levels of dominance, the number of competitors is a preponderant factor in determining the frequency of the convergent associate, but for higher levels, the absolute frequency of the most and least dominant item may be the more dominant factor.

Subjects

Summer school children of approximately nine and ten years of age who had completed third grade were randomly assigned to four independent groups. At the time of testing they were in attendance at the San Leandro Summer School which is conducted by the San Leandro Unified School District. In no case was the program they were taking remedial.

Experimental Design and Materials

Four different 18 pair lists were employed in which 18 different stimulus words were paired with nine different response words. East list was presented in three different random orders with the same starting order given to each S. Stimuli paired with the same response were never presented consecutively. Independent groups received each list.

The experimental design consisted basically of an independent groups design in which two list characteristics were varied as a two-by-two factorial design. The two characteristics were the extent of the variance among the stimulus pairs and the strength of the convergent primary. In addition, three levels of convergent associative strength of connection between each stimulus pair and its response-to-be-learned was varied within lists as a repeated measure.

The four lists which were employed appear in Table 1 of the Appendix. The characteristics which describe the doublets within each subgroup, for each list are shown in Table 1. It was intended that they vary symmetrically and systematically in the respects mentioned above, be as constant as possible with respect to the strength of the primary responses to the individual words and in the interitem associative strengths, but this ideal was only approximated because of the limited number of pairs which were available. Responses of fourth graders from the Palermo-Jenkins norms were used.

Pilot Study

Before running the Ss listed above, the lists were given preliminary test on 4th grade Ss in the same district, but this preliminary work on approximately 20 Ss indicated that these materials were too easy, and thereupon, four younger children were tested as the basis for using the younger age group for the present experiment.

Experimental Procedure

Standard paired-associate learning with a 2-2-2-rate and a 6 sec. intertrial interval was employed using a Phipps and Bird memory drum. After E obtained the S's name and age, the standard instructions were given, as follows:

"This is a task to see how well you can remember words. Here on this side (point to left aperture) you will see a word. A few seconds later the window on this side (point to right aperture) will open and you will see a second word. Your job is to try to tell me what the second word will be before the second window opens. There are eighteen pairs of words and the same two words go together every time. After the eighteen pairs there will be a blank space each time.

So that you will know what the pairs are going to be, the first time we go through the list I want you to just read both words in each pair out loud. Then after that I want you to begin guessing what the second word will be before you see it. You do not need to say the first word every time. In the beginning you will make a lot of mistakes, but after you learn how it goes you won't make so many. Just remember, try to say the second word out loud before the second window opens.

Is that clear? Alright, let's start."

After a familiarization trial, during the 6- sec. between trial interval: "Very good. Now begin guessing the second word before you see it."

One familiarization (study) trial preceded the anticipation trials. All Ss were given ten trials.

Results

Number of Correct Responses and Overt Errors. Analyses of variance were conducted with two measures of rate of learning, number of correct responses and overt errors. They are presented together since they are highly similar in nature and yielded substantially the same results. In addition, the analysis of each dependent variable was conducted twice. For the one, the doublets within each list were divided into three sets, high, medium, and low, on the basis of convergent associative strength; for the other, the same doublets were similarly divided into three sets on the basis of the mean dominance of the response from the single-word norms.

The means for the four lists appear in Table 2. The main effects of Strength of Convergent Primary and Variance were not significant in any of the analyses in question. However, the effect of Variance showed a trend in both analyses of number of correct responses ($F = 3.48$ and 3.44, d.f. 1/60, $p < .10$); i.e., more correct responses were given to high variance doublets than to low variance doublets for the ten trials as a whole. However, although the low variance doublets were learned more slowly, more errors were made on the high variance doublets initially. Table 3 and Figure 1 illustrate this interaction of Variance and Trials which was reliable for one analysis of overt errors (classification of pairs by convergent strength) and showed a strong trend for the other ($F = 2.89$, d.f. 4/840, $p < .10$ and $F = 3.63$, d.f. 4/840, $p < .025$).

For most lists, analysis of the doublets grouped by convergent associative strength entailed many shifts in items from one category to another, as compared to those employed in the analyses based on mean associative dominance. However, the results were highly similar nonetheless. The main effects of associative strength, shown in Table 4, were significant in all relevant analyses, $p < .005$.[1]

These interactions also indicated that high strength items are more difficult to learn when their primaries are weak than when they are strong, but that low strength items are more difficult when their primaries are strong. In general, the low strength items by either definition were more difficult than the high strength items, but the differences between the high and medium strength items were not consistently as anticipated. However, for both types of associative strength the hypothesized order of difficulty obtained for analysis of number of correct responses. The interactions with trials were significant and highly similar in the analysis of convergent and mean associative strength as shown in figures 2[2] and 3 and Table 5 ($F = 4.80$ and 3.62, d.f. 8/840, $p < .005$, respectively). The same interaction of Associative Strength and Trials was also shown for number of overt errors when associative strength was measured by mean dominance ($F = 2.54$; d.f. 8/840; $p < .025$), but did not show a comparable trend for convergent associative strength. Table 6 presents the relevant means. The error data for both types of associative strength indicates a significant interaction between Strength of Primary, Associative Strength, and Trials. Figures 4 and 5 illustrate the consistency for the pairs having a strong convergent primary, between the error data and the results found with number correct. However, for pairs having a weak convergent primary, there were deviations which occurred for the first two trials. The most marked of these was that for analyses of mean dominance the high mean dominant pairs were the most difficult for the first two trials, and for the analyses based on convergent strength, the medium convergent strength pairs were most difficult within those lists, for the first two trials. Thereafter, the data became increasingly more consistent with the expected pattern: That difficulty would decrease with strength.

The interaction between both types of Associative Strength and Strength of Convergent Primary was significant $p < .005$ in all four relevant analyses. The means appear in Tables 7 and 8, and indicate that the expected decrease in difficulty with increased associative strength does occur for the pairs having strong convergent primary, but does not show as strong or as consistent effect for the pairs having weak primaries. However, in all cases the low associative strength pairs are the most difficult.

The interaction between Variance and Associative Strength was significant in all relevant analyses.[3]

[1] For the analysis of Convergent Strength, the F values for number correct and number of overt errors were, respectively, $F = 78.76$ and 31.82, d.f. 2/840. For the analysis of Mean Dominance, the corresponding values were $F = 40.88$ and 14.57.

[2] The linear trend for the interaction shown in Figure 2 was significant, ($p < .005$) but the quadratic trend was not.

[3] The interaction of Associative Strength and Variance are described in the following sequence, number correct followed by number of overt errors for convergent associative strength, ($F = 31.82$ and 17.89, d.f. 2/840, $p < .005$) and number correct followed by number of overt errors for mean associative strength ($F = 47.50$ and 22.92, d.f. 2/840, $p < .005$).

The progression in difficulty from the low to high strength pairs was consistently more marked in the low variance pairs. (Tables 9 and 10 show the relevant means.) However, the progression in difficulty from the low to the high strength pairs depended not only on variance but also on Strength of Primary, as illustrated by the significant triple interaction among these variables.[1] Specifically, the means shown in Table 1 indicate that the progression in difficulty from the low to high strength pairs show small departures from regularity (reversals involving medium strength items), with the exception of one set of means which is regular and one which entails a large departure from regularity (fewer correct responses to low strength pairs than to high). The latter effect occurred in the high variance pairs having weak convergent primaries.

The error data, shown in Table 11, show a consistent decline in difficulty from the low strength to the high strength pairs which is more marked for low variance pairs and for pairs having strong convergent primaries, with the discrepancy though minimal, again appearing in pairs for which the convergent primary was weak and the variance was high.

Grouping Scores. Scores were computed for each S for each group of doublets divided by means of convergent associative strength. The method used was devised by Chalmers and entails the computation of the chance frequency with which errors would occur in response to both pairs within a doublet, given every possible number of errors on that trial. For instance, if two errors occur on one trial, the chance probability that both would occur in response to pairs within a doublet is 2/17. The expected clustering score (number of errors occurring in response to the pairs within doublets) is given by the formula:

$$C = \frac{n\ (n-1)}{N - 1}$$

where n is the number correct on a given trial and N is the number of pairs in the list. The formula pertains to doublets although a similar version could be developed for larger sets of related pairs.

The mean differences between the actual number of errors which appeared within couplets and the number which would be expected by chance was obtained for each trial and for each S. The Ss' average difference was used as the dependent variable in an analysis of variance in which the convergent associative strength was the repeated measure and Variance and Strength of Primary were the experimental variables. The effect of Variance was statistically significant ($F = 7.44$, d.f. 1/60, $p < .025$), Convergent Strength was not a factor ($F > 1$), and Strength of Primary showed a trend ($F = 2.94$, d.f. 1/60, $p < .10$). Specifically, larger degrees of clustering occurred in response to low variance pairs than to high, the means being .21 and -.03, respectively, and some tendency was observed for more clustering to occur in response to pairs having strong primaries, the means being .16 and .01, respectively. Although there was no evidence of any main effect of convergent strength, there was a significant triple interaction ($F = 2.10$, d.f. 2/120, $p < .005$). The relevant means appear in Table 12. The table suggests that the difference between low and high variance pairs occurred only for the medium and high strength pairs having a high convergent primary and for the low strength pairs having a low convergent primary.

It may be of interest to compare Table 11 with Table 12 in order to assess the relationship between performance and clustering. It would appear that for the relatively easy pairs (high and medium in strength), the tendency to cluster is not clearly related to performance, but for low variance pairs having weak primaries the tendency to cluster is associated with poor performance.

Error Analyses. The data for each S were processed with respect to the frequency of four types of overt errors: The total number of stimulus intrusions (S-errors); the total number of stimulus intrusions which consisted of substituting for the correct response the other stimulus within the same doublet (G); the total number of response intrusions (R-errors); the total number of extralist intrusions (E-errors). Table 13 presents the total frequency of these errors for each subgroup. The total number of S errors includes the G errors. On a chance basis only 1/17 of the S errors would be G errors while all of the obtained percentages were greater than 1/3. This measure concerned only overt stimulus errors, in contrast to the clustering measure which employed all types of overt errors, but only made use of errors which occurred on those trials on which 2, 3, or 4 errors occurred in response to each subgroup of three doublets. Also, the present measure, G errors or the proportion G/S (%), clearly indicates grouping. A comparison between the two measures is afforded in Table 14 in which the proportion of G-errors from among the total number of S errors is presented together with the cluster scores. For both measures, less evidence of grouping exists for the weak primary, high variance list than the others. High agreement between the % G errors (the proportion of the total number of overt errors which are G errors) and the cluster scores as shown in the last two columns of Table 13. The condition which showed maximum evidence of grouping was the strong primary, low variance condition and the other low variance condition was the second greatest in this respect.

---

[1] The interaction of Associative Strength, Variance and Strength of Primary were, in the order indicated above, ($F = 17.43$ and 6.58, d.f. 2/840, $p < .005$) and ($F = 7.72$ and 2.35, d.f. 2/840, $p < .005$).

An analysis of variance was carried out with G-errors expressed as a percentage of the total number of overt errors made by each S. Despite the possible inhomogeneity in the scores, and the lack of normality which would be assumed by the use of percentages, a constant was added for the purposes of the analysis and no transformations were made. The main effects of Strength of Convergent Primary and Variance were significant (F = 5.15; d.f. 1/60; $p < .05$ and F = 5.41; d.f. 1/60; $p < .025$), respectively. The means shown in the final table indicate that a significantly greater proportion of overt errors consisted of the other stimulus within the doublet for those lists containing doublets with strong rather than weak primaries and low variance rather than high variance doublets. In other words, the evidence suggests greater cohesiveness of low variance doublets and those with strong convergent primaries.

Discussion

The results indicate large effects of associative strength and interrelationships between this variable and the other major experimental variables. However, the effects of mean dominance and convergent strength are not differentiable on the basis of the analyses employed. In general, low strength pairs were more difficult to learn than high strength pairs, and the means deviated from this pattern in only one case, for high variance pairs having weak primaries. In this case the performance on the low strength pairs was markedly better than performance on the low strength pairs in the other three lists. It was also true that there was significantly less clustering in this deviant list than in any other list; the amount of clustering employed being definitely less than chance. This suggests that these Ss were deliberately avoiding clustering - and it helped them learn!

As might be expected, the differences in difficulty among the high, medium, and low strength doublets diminished over trials. Moreover, the differences among the high, medium, and low strength items were more marked and regular for the items having strong convergent primaries and for which the variance was low. This was due in part to the significant superiority of high strength items having strong primaries over high strength items whose primaries were low. This effect is partially analogous to the effect obtained with strength of single-word primary, which suggests that it is easier for children to learn new items in response to words having strong primaries than to words having weak ones (Wicklund, 1964). In this case, it was easier to learn responses to doublets having strong convergent primaries. One would expect that more clustering would have occurred in the case of doublets with strong convergent primaries, but this was not consistently the case. For relatively strong items having strong primaries, clustering only occurred when the variance was low. Thus it would seem that Ss did not cluster through the mediated primary response, but rather in response to the particular response-to-be-learned.

The fact that more clustering occurred in response to low variance pairs than to high variance pairs could have reflected an artifact: that the low variance pairs within a doublet were more nearly equal in difficulty than any other pairs within the list. Thus, for high variance lists, pairs outside of doublets could have been more nearly equal in difficulty and this would reduce the observed clustering therein. This possible artifact does not, however, explain the fact that as hypothesized, there tended to be more clustering in doublets for which the convergent primary was strong. Nor does it affect the fact that for low strength items there was less clustering in high variance doublets when the strength of the convergent primary was weak than when it was strong. Moreover, it could not concern the corroboration of the cluster score findings by the relative frequency of G-errors.

The reason for varying Strength of Primary and Variance was that it was hypothesized to affect grouping during learning, which was assumed to be facilitative. There is no doubt but that these variables did affect grouping in the hypothesized direction although the effects were less consistent than would be desired. However, the expected facilitative effect of grouping is clearly not upheld, and there is a suggestion that under some circumstances it may actually interfere.

Although the high variance pairs tended on the whole, to be learned more readily than low variance pairs, or the first two trials the effect was reversed. At the outset, learning would be expected to reflect more directly the associative strengths between the members of the pair than later learning, and indeed, the effect was consistent with the expectations based on the frequencies of the convergent response as a function of mean dominance and variance, wherein, within the ranges investigated, and where reliable results obtained, the incidence of the convergent response to low variance pairs was greater than in response to high variance pairs. Although the reasons for a shift in the relative difficulty of low and high variance pairs may not be immediately obvious, they can be explained ad hoc. Specifically, it seems likely that in high variance doublets, the more dominant pair would be learned very quickly, leaving only the less dominant pair to be learned. For the low variance doublets, both pairs would take longer to be learned than the more dominant pair from the high variance doublet. Thus, for high variance doublets, the effective length of the list would be halved for most of the learning period, facilitating the Ss ability to learn the remaining half. For the low variance doublets, the effective length of the list would thus remain relatively long for mor trials during learning, thus creating greater difficulty in learning the pairs.

Difficulty of learning within lists could be, in part, explicable by the imperfect matching of items in the relevant characteristics. However, the results do not, on the whole, support this interpretation. For example, the fact that for medium and low strength doublets the overall result obtained that high variance doublets were easier to learn than low variance doublets, but for high variance doublets the trend was in the opposite direction. Inspection of initial strengths reveals that in most cases, the initial strengths of the low variance doublets were higher; yet, they were actually found to be harder to learn than the high variance doublets. The obtained result is thus consistent with the shortened list interpretation described above. For the high strength pairs, any differences in difficulty between high and low variance lists would have to appear relatively early in learning during which time performance on the high variance lists would be expected to be superior. Since all items under both variance conditions would then be learned quite readily, the difference in number of trials to learn the two pairs of a doublet would be relatively small and therefore, the advantage of shortened list length for the high variance pairs would be reduced.

The tendency to group pairs within doublets during learning was hypothesized to be greatest for doublets having strong convergent primaries and low variance, and in fact, clear confirmation of these hypotheses was obtained by two independent analyses: the frequency of G-errors and the degree of coupleting relative to chance values. For the latter analysis, the effect of strength of primary did not quite reach statistical significance ($\underline{p} < .10$) but the means were in the expected direction. Also, for the latter analysis, the effect of grouping was found to depend on the convergent associative strength of the doublets. (This variable could not be analyzed for the other measure.) In general, there was little evidence for grouping any pairs within doublets when the variance was high, but when the variance was low, the possibility of such grouping was strengthened. Specifically, for doublets of low variance, strong grouping occurred among the high and medium strength items having strong convergent primaries. However, some grouping also occurred in low strength doublets having weak primaries. The effects cannot be explained in terms of the frequency with which the convergent response-to-be-learned was also the convergent primary. This did occur in seven cases, all of them in high strength doublets, but none in the set of doublets characterized on the one hand, by low variance and strong convergent primaries and on the other hand, by having been grouped during learning. The major effect seems to be that low variance promotes grouping, and the extent to which it occurs depends on the strength of the response-to-be-trained as a mediator and on the strengths of other potential mediators, for which the convergent primary is a prime candidate. One way in which low variance promotes grouping concerns the fact that the pairs within low variance doublets are more equal in difficulty than pairs within high variance doublets; and this equality entails a longer period prior to learning during which such grouping could occur. Also, any linkages which occur after learning of one or both pairs within a doublet would be harder to detect. They would not appear on the coupleting measure and the likelihood of G-errors would be drastically reduced after the learning of one pair through the stimulus and response differentiation which inevitably accompanies learning.

In general, the results concerning associative strength are highly similar for the two types. However, if pair difficulty rather than doublet difficulty were analyzed as the unit, it is very likely that associative dominance would relate more directly to rate of acquisition than would convergent strength.

TABLE 1

Mean Values (%) for Each Subgroup Within Each List of the Manipulated and other Stimulus Characteristics, Showing *also* the Mean Number of Correct Responses During Learning

Dominance	Mean Dominance	Convergent Strength	Variance	Strength of Convergent Primary	Best Single	Combined Probability	Mean Number Correct*
ASSOCIATIVE STRENGTH							
Weak Convergent Primary							
Low Variance							
Low	1.37	8.73	0.07	14.68	1.40	8.67	8.91
Med	7.20	7.01	3.73	10.05	9.07	13.83	8.71
High	16.50	18.39	10.20	23.28	22.00	30.57	9.72
Overall	8.36	11.38	4.67	16.00	10.82	17.69	9.1
High Variance							
Low	1.57	2.91	0.73	13.23	1.93	3.13	10.14
Med	4.50	8.99	4.87	9.92	6.93	8.87	10.61
High	15.80	15.34	22.93	15.87	27.27	30.40	9.92
Overall	7.29	9.08	9.51	13.00	12.04	14.13	10.2
Strong Convergent Primary							
Low Variance							
Low	1.87	3.97	0.93	27.64	2.33	3.70	8.79
Med	5.23	5.29	2.87	24.20	6.67	10.23	8.40
High	18.00	13.62	6.13	24.74	21.07	32.27	11.00
Overall	8.37	7.63	3.31	25.53	10.02	15.40	9.1
High Variance							
Low	2.53	5.55	2.27	24.07	3.67	5.00	8.64
Med	12.90	5.56	21.07	27.19	23.73	25.67	10.56
High	22.43	20.50	35.47	20.50	39.80	43.20	10.14
Overall	12.62	10.54	19.60	23.92	22.40	24.62	9.8
CONVERGENT ASSOCIATIVE STRENGTH							
Weak Convergent Primary							
Low Variance							
Low	3.69	3.37	1.27	12.04	4.00	12.50	7.75
Med	7.66	8.50	5.13	12.57	11.13	16.30	9.90

Continued on next page.

2

Table 1, continued
Convergent Associative Strength

High	22.74	13.20	7.60	23.41	17.33	24.27	9.80
Overall	4.66	8.36	4.67	16.00	10.82	17.69	9.1
			High Variance				
Low	2.90	1.57	0.73	13.23	1.93	3.13	10.14
Med	8.99	5.77	7.27	10.98	9.40	11.33	10.61
High	15.35	14.53	20.53	14.81	24.80	27.93	9.92
Overall	9.08	7.29	9.51	13.00	12.04	14.13	10.2
			Strong Convergent Primary				
			Low Variance				
Low	3.18	2.23	0.93	28.97	2.33	4.43	8.40
Med	6.07	4.87	2.87	22.88	6.67	9.50	8.79
High	13.61	18.00	6.13	24.74	21.07	32.27	11.00
Overall	3.31	8.37	3.31	25.53	10.02	15.40	9.1
			High Variance				
Low	2.51	3.67	4.67	23.88	3.67	7.23	8.84
Med	8.60	11.77	18.67	27.38	23.73	23.43	10.39
High	20.50	22.43	35.47	20.50	39.80	43.20	10.14
Overall	19.60	12.62	19.60	23.92	22.40	24.62	9.8

* Mean number correct per block of two trials, for six pairs: maximum correct = 12.

TABLE 2

Mean Number of Correct Responses, Overt Errors, and Grouping Scores in the Four Lists[1]

(Showing the Main Effects of Strength of Primary and Variance, and Their Interaction)

Strength of Primary

Variance		High	Low	Overall
			Correct Responses	
	Low	9.40	9.16	9.28
	High	9.78	10.23	10.00
	Mean	9.69	9.59	
			Overt Errors	
	Low	.57	.75	.66
	High	.55	.52	.54
	Mean	.56	.63	
			Grouping Scores	
	Low	.31	.11	.21

Continued on next page.

Table 2, continued
Grouping Scores

High	.02	-.08	-.03
Mean	.16	.01	

---

[1] Means refer to number of items, errors, scores per subgroup of 6 pairs for each block of two trials.

TABLE 3

The Mean Number of Overt Errors Given to High and Low Variance Doublets for Each Block of Two Trials

		Trial				
		1-2	3-4	5-6	7-8	9-10
Variance						
	Low	.86	.84	.69	.51	.41
	High	1.09	.69	.43	.39	.05
	Overall	.98	.76	.56	.45	.23

TABLE 4

The Effect of the Two Types of Associative Strength on Number of Correct Responses and Overt Errors

	High	Medium	Low
Convergent Associative Strength			
Number Correct	10.22	9.92	8.79
Number Overt Errors	.39	.51	.88
Mean Dominance			
Number Correct	10.20	9.60	9.12
Number Overt Errors	.43	.57	.79

TABLE 5

Mean Number of Correct Responses in Each Block of Two Trials for Doublets of Varying Associative Strength

	Blocks of Trials				
	1-2	3-4	5-6	7-8	9-10
Convergent Associative Strength					
High	8.13	10.03	10.80	10.88	11.25
Medium	7.67	9.69	10.38	10.63	11.25
Low	5.67	8.06	9.56	10.03	10.63

Continued on next page.

Table 5, continued

Mean Associative Dominance

High	8.13	10.09	10.77	10.80	11.20
Medium	7.19	9.17	10.14	10.41	11.11
Low	6.16	8.47	9.83	10.33	10.83

TABLE 6

Mean Number of Overt Errors Over Blocks of Two Trials as a Function of Mean Associative Dominance

	Trials					
Associative Dominance	1-2	3-4	5-6	7-8	9-10	Overall
High	.98	.47	.30	.33	.11	.43
Medium	.78	.65	.63	.45	.34	.57
Low	1.20	1.22	.73	.55	.25	.79
Mean	.99	.78	.55	.43	.23	

TABLE 7

Number of Correct Responses Within Blocks of Two Trials as a Function of the Two Types of Associative Strength and Strength of the Convergent Primary

Convergent Associative Strength

Strength of Primary	High	Medium	Low	Overall
Strong	10.57	9.59	8.61	9.59
Weak	9.86	10.26	8.97	9.69
Overall	10.22	9.92	8.79	

Mean Associative Dominance

Strength of Primary	High	Medium	Low	Overall
Strong	10.57	9.48	8.72	9.59
Weak	9.82	9.73	9.53	9.69
Overall	10.20	9.60	9.12	

TABLE 8

Number of Overt Errors Within Blocks of Two Trials as a Function of the Two Types of Associative Strength and Strength of the Convergent Primary

Convergent Associative Strength

Strength of Primary	High	Medium	Low	Overall
Strong	.26	.54	.86	.56
Weak	.53	.48	.90	.64
Overall	.39	.51	.88	

Mean Associative Dominance

Strength of Primary	High	Medium	Low	Overall
Strong	.26	.58	.85	.56
Weak	.60	.57	.73	.63
Overall	.43	.57	.79	

TABLE 9

Mean Number of Overt Errors per Block of Two Trials Given to Doublets Varying in Two Types of Associative Strength and Variance

	Convergent Associative Strength			
Variance	High	Medium	Low	Overall
Low	.27	.77	.95	.66
High	.52	.26	.81	.54
Overall	.39	.51	.88	
	Mean Associative Dominance			
Low	.34	.89	.74	.66
High	.52	.25	.84	.54
Overall	.43	.57	.79	

TABLE 10

Mean Number of Correct Responses in Each Block of Two Trials Given to Doublets Varying in Two Types of Associative Strength and Variance

	Convergent Associative Strength			
Variance	High	Medium	Low	Overall
Low	10.40	9.34	8.09	9.28
High	10.03	10.50	9.49	10.00
Overall	10.22	9.92	8.79	
	Mean Associative Dominance			
Low	10.36	8.62	8.85	9.28
High	10.03	10.59	9.39	10.00
Overall	10.20	9.60	9.12	

TABLE 11

The Effect on the Mean Number of Overt Errors per Block of Two Trials of the Strength of the Convergent Primary on the Relationship Between Convergent Associative Strength and Variance

	Convergent Associative Strength		
Variance	High	Medium	Low
	Strong Convergent Primary		
Low	.15	.82	.75
High	.38	.26	1.01
	Weak Convergent Primary		
Low	.39	.71	1.15
High	.66	.50	.65

TABLE 12

Mean Clustering Scores as a Function of the Major Experimental Variables

	Strength of Convergent Primary			
	High		Low	
	Low Variance	High Variance	Low Variance	High Variance
Convergent Strength				
High	.36	.02	.07	-.12
Medium	.59	-.09	-.05	.08
Low	-.02	.13	.30	-.19

TABLE 13

Total Frequency of Each Type of Error Within Each List

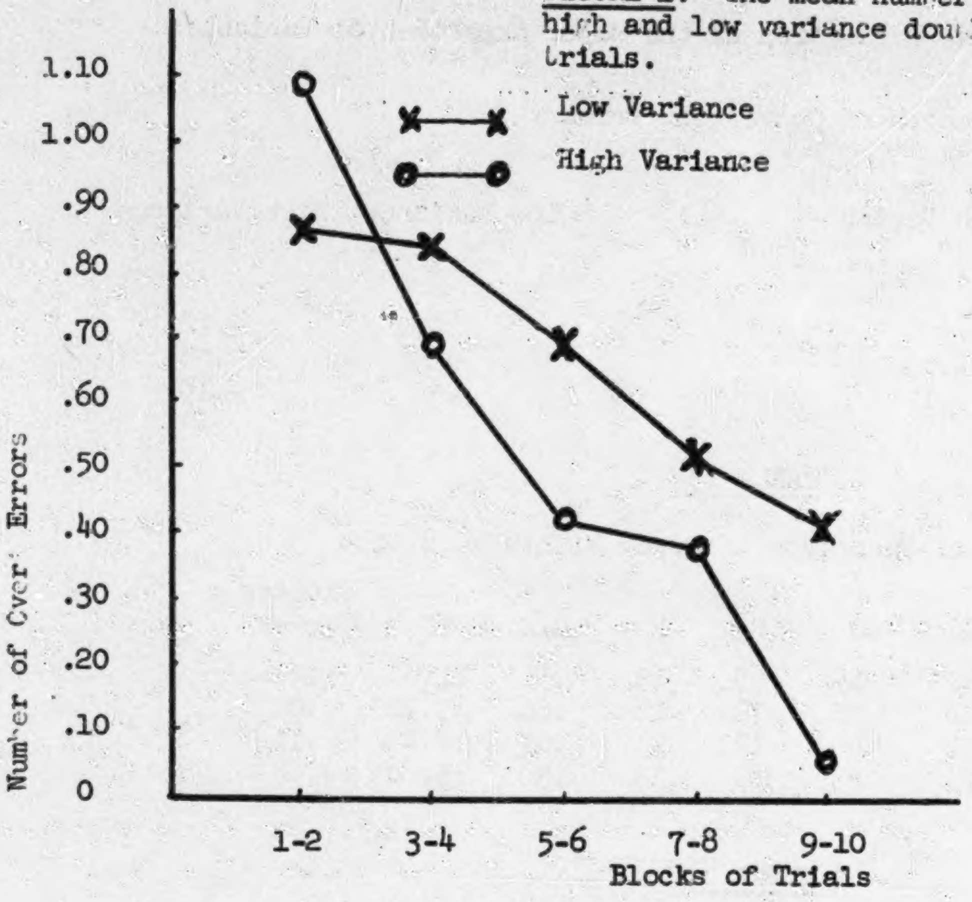
	(G)	S	R	E	G/S	%G	Cluster Scores
Strong Primary, Low Variance	30	59	65	13	.51	21.90	.31
Strong Primary, High Variance	21	32	72	28	.66	15.91	.02
Weak Primary, Low Variance	31	46	123	9	.67	17.42	.11
Weak Primary, High Variance	7	23	89	13	.30	5.60	-.08

TABLE 14

The Frequency of G-Errors (%) in the Four Lists, Showing the Main Effects of Strength of Primary and Variance

	Strength of Primary		
Variance	Weak	Strong	Mean
Low	14.93	29.81	22.37
High	7.01	14.65	10.83
Mean	10.97	22.23	

FIGURE 1. The mean number of overt errors given to high and low variance doublets for each block of two trials.



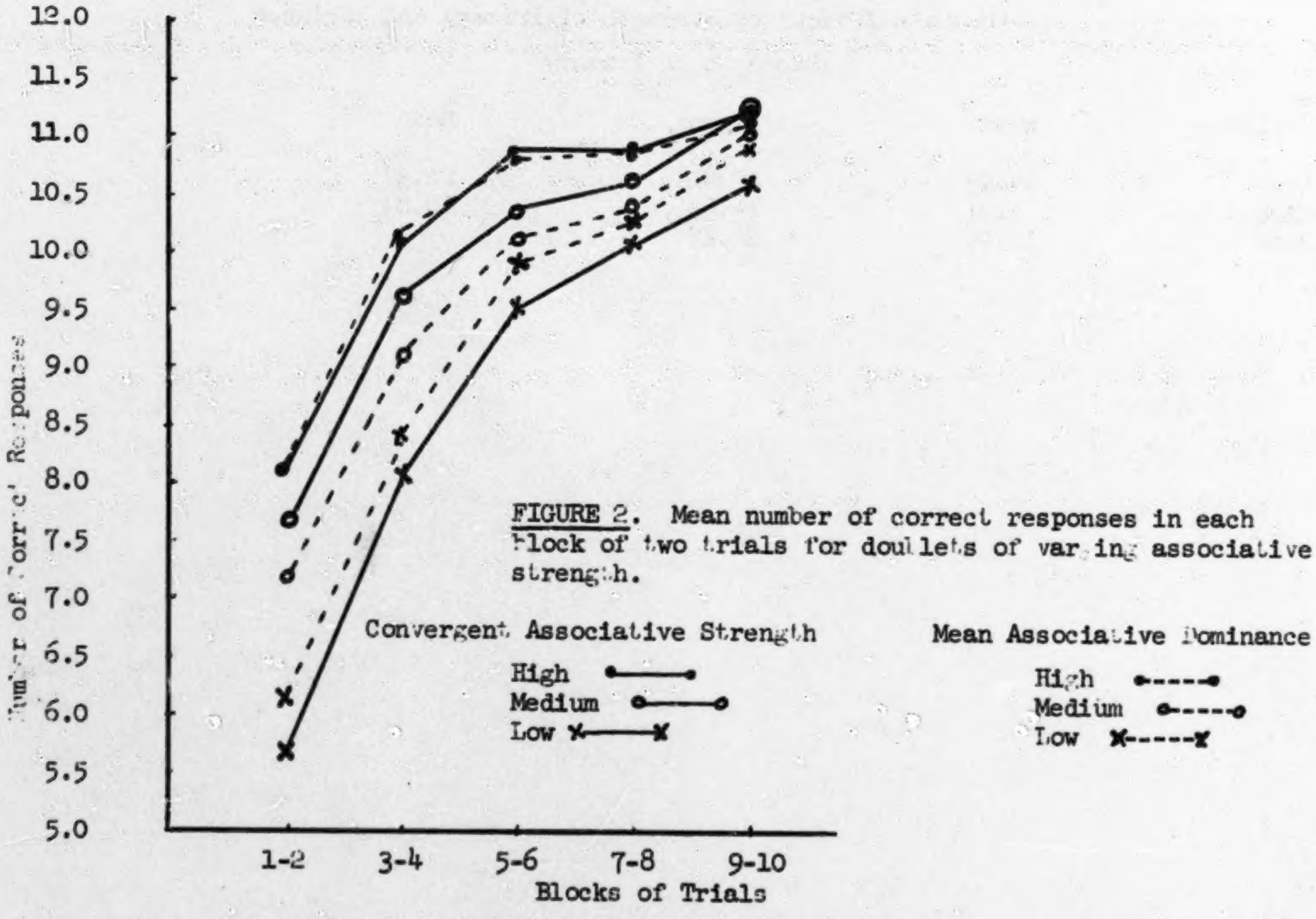


FIGURE 2. Mean number of correct responses in each block of two trials for doublets of varying associative strength.

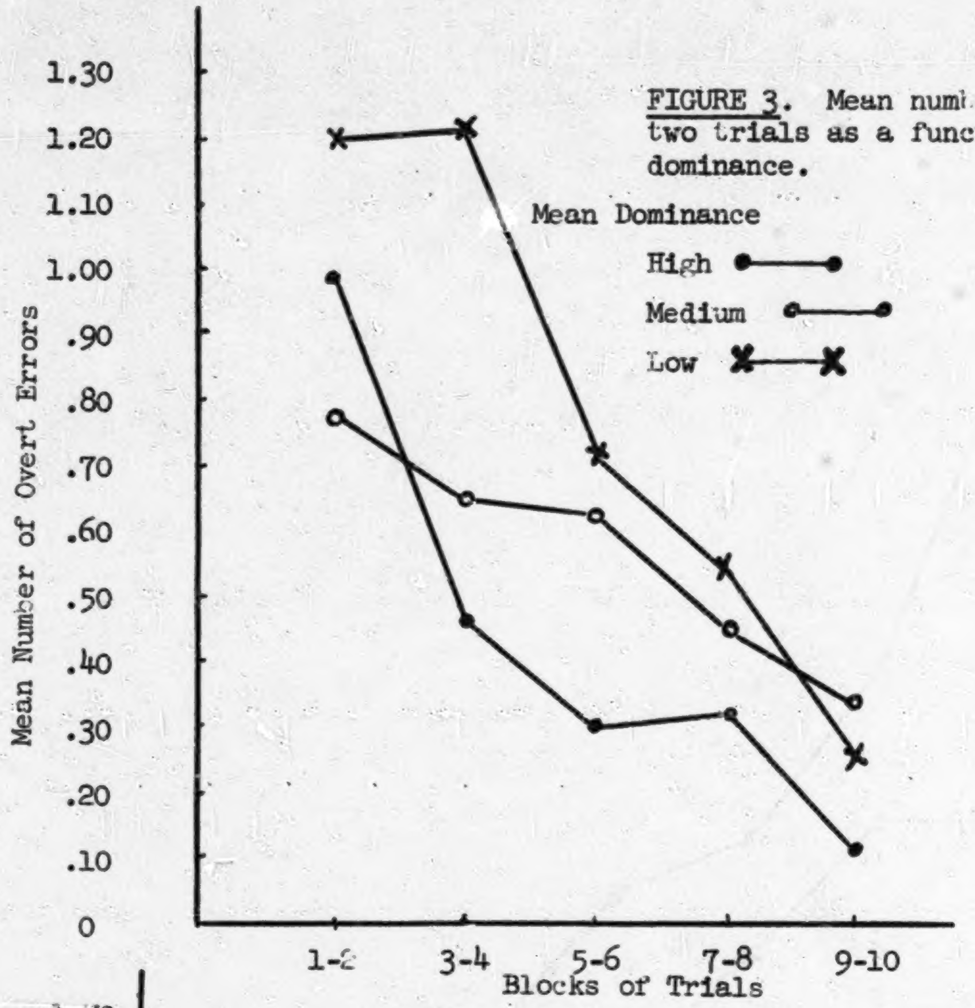


FIGURE 3. Mean number of overt errors over blocks of two trials as a function of strength of mean associative dominance.

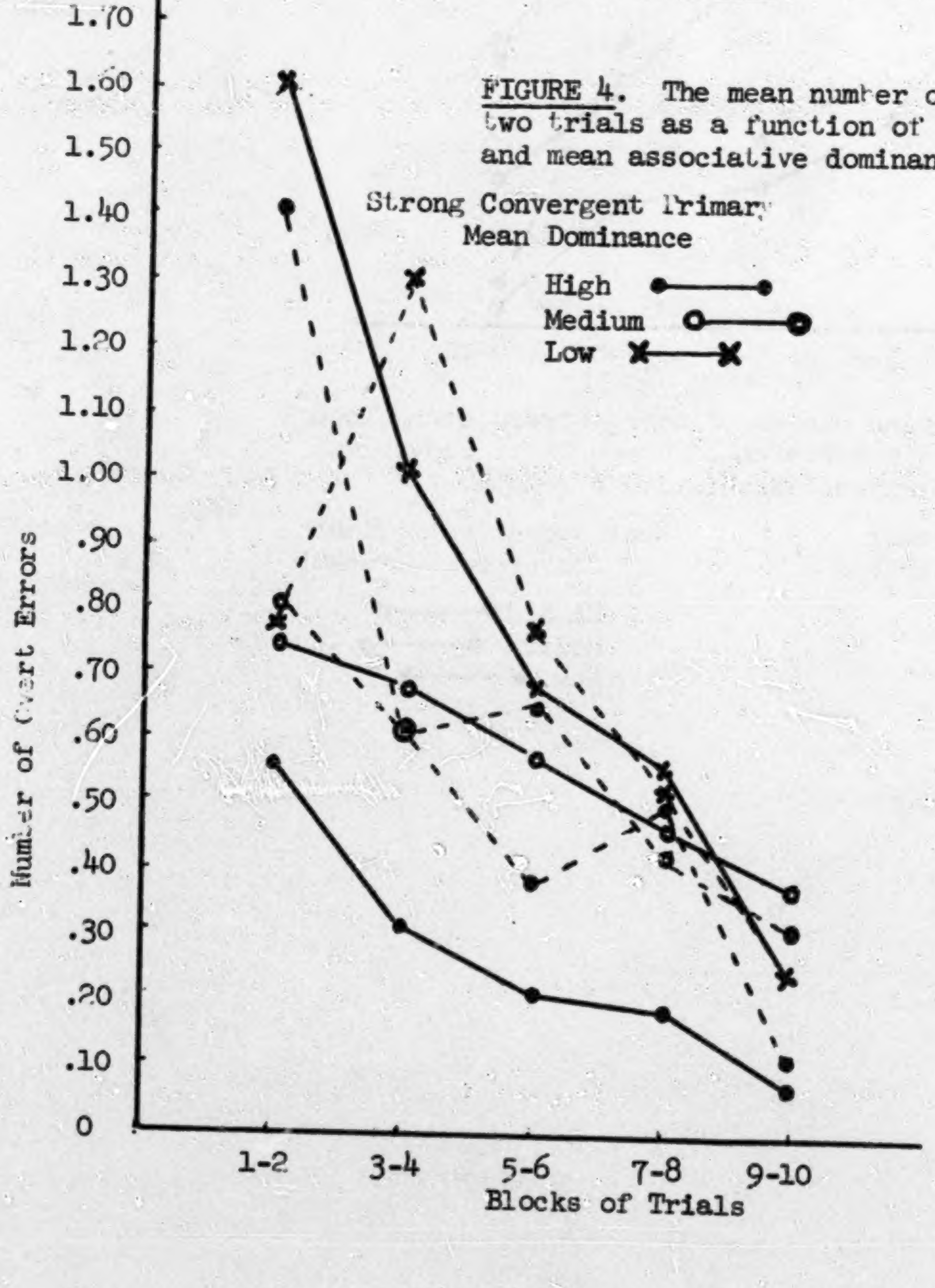


FIGURE 4. The mean number of overt errors over blocks of two trials as a function of the strength of convergent primary and mean associative dominance.

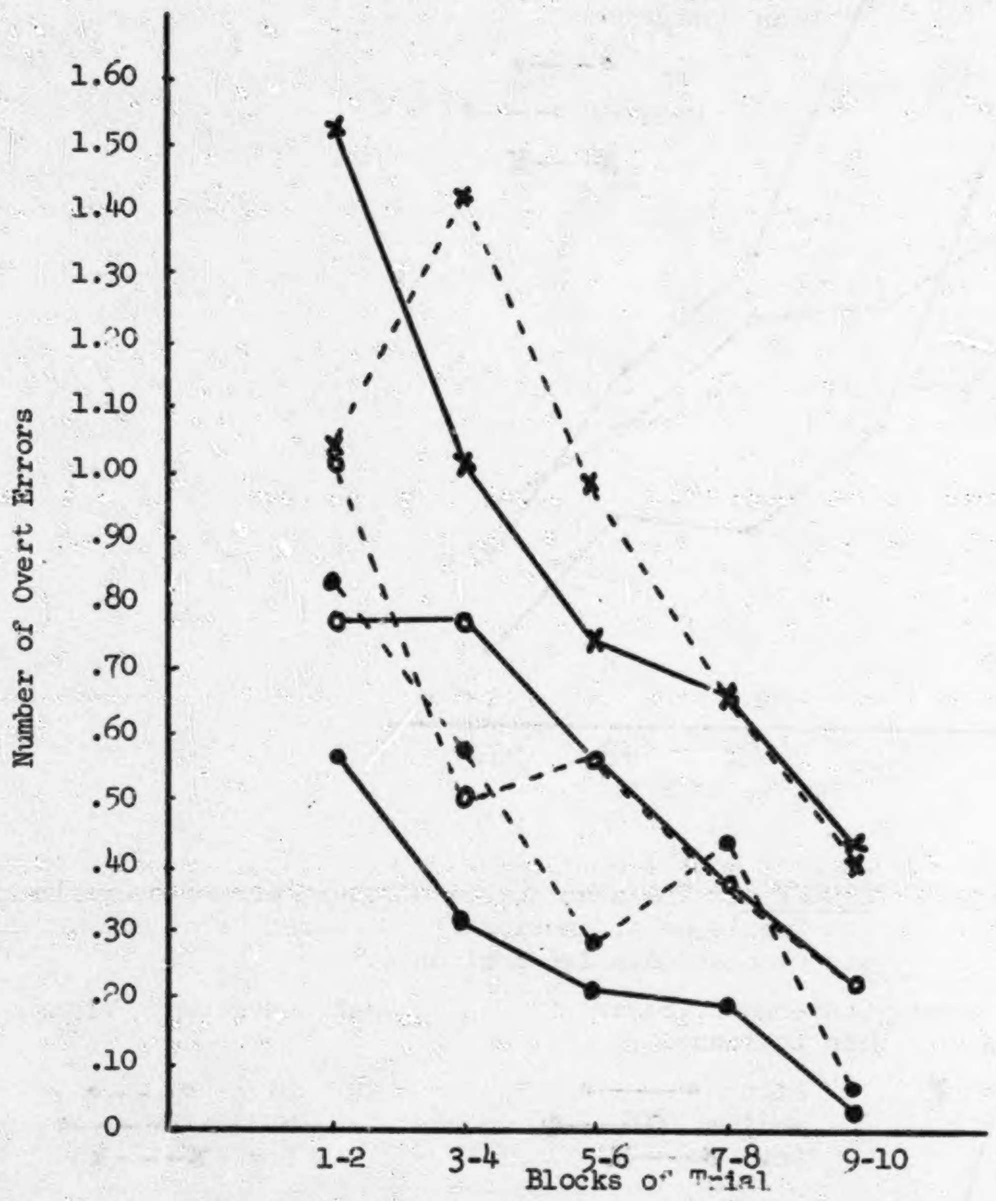


FIGURE 5. The mean number of overt errors over blocks of two trials as a function of strength of convergent Primary and convergent associative strength.

Strong Convergent Primary	Weak Convergent Primary
Convergent Strength	Convergent Strength
High ●——●	High ●——●
Medium ○——○	Medium ○——○
Low ✖——✖	Low ✖——✖

Chapter 7

Learning Ability of Children as a Function of Type of Mediating Relationship[1]

It is known that the tendency of children to group words in conceptual categories increases with age (Bousfield, Esterson and Whitmarsh, 1958). The tendency for children to group words as a function of their associative properties was investigated in the experiment on paired-associate learning reported in chapter 6. In that case, the stimuli of the pairs comprising a doublet were words having a known association in common. However, adults create associates to words, even when the words are thought to be unrelated on the basis of associative norms (Tulving, 1962) and it is reasonable that children of increasing age and ability should show an increasing tendency to link verbal items on a conceptual basis. Consider pairs which form doublets by virtue of the fact that a common response is to be learned to the two stimuli. If the stimuli are linked by the *S* in some conceptual manner, he should learn the doublet as a unit; i.e., learn the response to both of the stimuli at approximately the same time. Learning of the items as doublets rather than as isolated units should also facilitate learning compared with rote learning of separate pairs as isolated units. If one pair is learned, the appearance of S_1 should elicit the conceptual link which would in turn elicit S_2. In this way S_1 and S_2 would be cognitively concurrent with R_{12} and this cognitive contiguity should facilitate the learning of $S_2 \longrightarrow R_{12}$. The pairs would be related in this way within a doublet:

$$\begin{tikzcd} S_1 \arrow[dr] & \\ & R_{12} \\ S_2 \arrow[ur] & \end{tikzcd}$$

If this conceptual linking has occurred, then it should be relatively easy to learn a new response to S_1 and S_2 when once again the same response is learned to the same two stimuli, as a doublet. However, if new responses are to be learned to stimuli in such a way that the doublets are destroyed and new doublets must be formed from the same set, interference relative to the previous condition in which the stimuli belonging to a doublet remain linked, would be anticipated. Since it was expected that older children would tend, more than younger ones, to link stimuli conceptually, more interference from such a transfer condition would be expected in older children.

In order to study this problem concerning the use of mediation in the learning of doublets, two variables, *Response Constancy* and *Stimulus Grouping*, were investigated with respect to transfer following the paired-associate learning of doublets. *Response interference* entailed re-pairing the same responses which had been learned during the acquisition stage and the *Response Control* condition entailed presenting new responses. The three conditions of *Stimulus Grouping* were Facilitation under which stimuli were maintained within the same doublet, Interference, under which stimuli had been present during the acquisition stage were re-grouped, and Control, under which completely new stimuli were supplied. Table 1 presents this experimental paradigm.

To study the role of mental ability and experience with respect to the spontaneous formation of conceptual groups during learning and with respect to the tendency to employ these groups in mediation, three groups of *Ss* were studied. First graders with a mental age of approximately six years, fifth graders with a mental age of about ten years, retardates whose chronological age was approximately the same as the fifth graders' and whose mental age was about the same as the first graders'. Thus, if mental age is the crucial factor, the first graders and the retardates should mediate similarly, but if years of experience with verbal materials is foremost in importance, the fifth graders and retardates should mediate in similar fashion.

---

[1] This experiment was designed and authored by Douglas Chalmers and Harriett Amster.

Method

Subjects

Seventy-two children from each of the following groups served as *Ss*: First graders, fifth graders, and retardates. They were selected so that the first graders and retardates might be assumed approximately equal in mental age while the fifth graders and retardates would be assumed equal in chronological age. Their respective mean C. A.s in years were 6.44, 10.44 and 10.60; their respective mean M. A.s were 6.44 (assumed), 11.45, and 7.32. The age ranges of the three groups were, respectively, six to seven years, ten to eleven years, and nine to thirteen years. Further information about subject characteristics appears in the Appendix. Intelligence test data were obtained for fifth graders and retardates, but the same tests could not be used consistently across all groups. However, the California Test of Mental Maturity was the instrument most frequently employed. IQs could not be obtained for the first graders and therefore, their mean mental age of the group as a whole was assumed to equal their mean chronological age. *Ss* were drawn from four schools in the San Leandro Unified School District in San Leandro, California.

Experimental Design

A factorial design involving independent groups was employed. The major experimental variables were *Response Constancy* and *Stimulus Grouping*. The two types of Response Constancy were response interference and response facilitation while the three types of Stimulus Grouping were facilitation, interference, and control. These conditions have been described in the Introduction and a specific table of stimuli actually presented to the *Ss* under each of these conditions appears in Table 2. In addition to the two major experimental variables, the experiment was replicated using a different but comparable set of materials, providing a third experimental variable, *Replication*. The three *Groups* of *Ss* comprised the fourth variable of the experiment. The respective design was a 2 x 3 x 2 x 3 factorial. The experiment entailed paired-associate learning of two successive lists. The experimental materials were selected in keeping with the preceding table, but the details of the initial pairings were worked out in accordance with the requirement that within replications, the same transfer (second) list be employed for *Ss* in all conditions.

Materials

The six item paired-associate lists described in the table comprised the materials. The stimuli were, in all cases, line drawings of the common objects named in the table. The responses were low meaningful bigrams for which the second letter followed the first with a mean associative frequency in percent < 1.5% for both second and fifth graders, using the Amster-Keppel (1966) norms. As shown in Table 2, in each list there were six different stimuli paired to three different responses, and this constituted three doublets. Sample lists, illustrating the line drawings are presented in the Appendix.

Pairs were presented randomly in four different orders with the restriction that no two pairs having the same response appear successively.

Procedure

Upon entering the room, each *S*, tested individually, was seated before a Stowe Memory Drum. *E* sat next to *S* throughout the experiment. Instructions included warm-up trials through a 2-item list of number pairs, presented on demonstration cards. Prior to testing, *Ss* were familiarized with the bigram responses which were typed on a card taped to the drum. They were thus readily available for reference throughout all stages of testing. All *Ss*, regardless of condition, saw the same card which consisted of 2 columns of 3 bigrams each: KG, RL, TC, and BX, FW, PJ. *Ss* were instructed that for any one task, only one of the colums of bigrams would be relevant.

During Task 1, it was necessary for *E* to interrupt the task for some 1st graders and retardates for various reasons. If an interruption occurred during the second task, *S*'s testing was terminated and he was replaced. Also during Task 1, *E* alternated uttering the phrases "You're doing fine" and "That's very good" at the end of each trial. The same procedure was employed with the phrases "Fine" and "Good" during the second task. Between the first and second tasks, there was a pause long enough for *E* to change tapes and to instruct *S* that his new task would be essentially the same as the first task.

The presentation rate throughout was 4:2 sec., with a 6 sec. intertrial interval.

Following the second task, all retarded children were administered a short form of the CAT.

Results

Task 1. All of the major hypotheses concerning the present experiment bear on the *Ss'* responses during the transfer task. For all major intents, the experimental treatments for the first paired associate task were equivalent for all groups. However, it is of interest to compare the initial learning performance of the three groups differing in mental and chronological age, and to examine any differences within these subgroups as indicative of errors of sampling or random variation in the difficulty of the materials. The number of trials to criterion was found to vary reliably among the three grade levels (*F* = 39.40, d.f. 2/180, *p* < .001). The first graders required 18.4 trials to criterion, the fifth graders 8.5, and the retardates 20.3. The Scheffé contrast indicated that the fifth graders learned reliably more readily (*p* < .01) than the first graders or retardates who did not differ from each other. The two replication groups were also found to differ reliably, (*F* = 9.50, d.f. 1/180, *p* < .01), but fortunately, there was no reliable evidence of any interaction between replication and any other variable. The mean number of trials to criterion for the *Ss* in replication A was 17.53 and for replication B was 13.94.

Transfer Task

The most crucial measure of the effect of the experimental treatments on proficiency consisted in the number of correct responses on the first transfer trial. Since the three grade levels and two replication groups differed in number of trials to acquire the original list, it is not surprising that they also differed in the number correct on the first transfer trial. Specifically, the interaction of Grade Level by Replication was significant (*F* = 3.43; d.f. 2/180; *p* <.05). The main effect of Grade Level was significant (*F* = 3.10; d.f. 2/180; *p*<.05) and the effect of Replication showed a trend (*p*<.10). The relevant means are shown in Table 3 which includes the means for original learning as a basis for comparison.

An analysis of the number of correct responses by blocks of two trials was carried out through trial 18. The effect of Grade Level was significant (*F* = 41.49, d.f. 2/180, *p* < .01). The mean number correct for the two-trial blocks was 7.33 for first graders, 9.49 for fifth graders and 7.02 for retardates. The effect of Stimulus Grouping was also reliable, (*F* = 16.79, d.f. 2/180, *p* < .01). The mean for the Facilitation condition was 7.46, for the Interference condition 7.44, and for the Control condition, 8.94. The Control condition was thus reliably more facilitative than the other two conditions which did not differ significantly from one another. Blocks of Trials was a significant variable indicating merely that the improvement due to practice was reliable, (*F* = 101.72; d.f. 8/1440; *p* < .001). Only the linear component was significant, *p* < .001. The interaction between Grade Level and Trials was significant, (*F* = 2.49; d.f. 16/1440; *p* < .05). The means are illustrated in Figure 1, showing that the rate of improvement is slowest for retardates and fastest for fifth graders. The Interaction between Stimulus Grouping and Trials was also significant, (*F* = 1.75; d.f. 16/1440; *p* < .05). These means are illustrated in Figure 2. Finally, the triple interaction shown in Table 4 of Response Constancy, Trials, and Replication was significant, (*F* = 2.79; d.f. 8/1440; *p* < .01). Figure 4 illustrates this interaction. No other effects involving Response Constancy were reliable.

In addition to considering the standard performance measures described above, a "clustering measure", developed by Chalmers (1965), was employed in order to determine whether or not *Ss* tended to learn the doublets as units to a greater extent than one would expect on a random basis. Clustering during transfer was assessed.[1] For one replication, 42% of the fifth graders cluster in the Facilitation condition while only 8% cluster in the Interference condition. The corresponding percentages for the first graders are 33% and 25%. On the basis of a sign test, considering the total number of fifth graders who cluster, a significantly greater number cluster in the Facilitation condition ($p < .01$) than in the interference condition. Considering the total number of first graders who cluster, no significant difference exists in the number who cluster in the Facilitation condition compared with the Interference condition. Considering the total number of first and fifth graders who cluster in the facilitation condition, no significant difference exists in the number of fifth graders who cluster, compared with the number of first graders who do so. However, the tendency for more first than fifth graders to cluster under the interference condition was clearly significant ($p < .01$). The actual amount of clustering which was accomplished overall tended to be at about the chance level, not above. Consequently, the fifth graders seemed to avoid clustering in the interference condition, rather than to actively cluster to a greater extent under the Facilitation condition.

---

[1] A more precise definition of this measure appears in the preceding chapter.

Discussion

Surprisingly, no facilitation in transfer due to maintaining stimulus groupings was observed for any of the three groups. The only facilitation came about from the presentation of entirely new stimuli, compared with maintaining the stimuli from the previous list. However, efficiency of performance did not, in this case, accurately reveal the mental processes employed by the various groups of Ss, whereas another dependent variable did. Specifically, the tendency to group the stimuli which belonged to a doublet was found to vary with both age and experimental condition, as hypothesized. Older children tended to group the stimuli during transfer.

The results indicated a clear developmental difference in tendency to cluster appropriately, such that normal children of ten years grouped conceptually to a greater extent than normals or retardates of a younger mental age. This finding is at variance with that of Osborn (1960) who found semantic clustering in recall to occur equally strongly in normal and retarded children. Quite possibly, it is the tendency to cluster when highly dominant conceptual categories are not readily available which distinguishes normals from retardates, but other factors e.g. task differences or differences in mental level may account for the results.

Table 1

Experimental Paradigm for Mediation Experiment I, Showing Experimental Design and the Conditions as Experienced[1]

	Stimulus Grouping					
	Facilitation		Interference		Control	
	Task I	Task II	Task I	Task II	Task I	Task II
Response Interference	$\begin{matrix}S_1\\S_2\end{matrix} > A$	$\begin{matrix}S_1\\S_2\end{matrix} > C$	$\begin{matrix}S_1\\S_2\end{matrix} > A$	$\begin{matrix}S_1\\S_3\end{matrix} > C$	$\begin{matrix}S_1\\S_2\end{matrix} > A$	$\begin{matrix}S_7\\S_8\end{matrix} > A$
	$\begin{matrix}S_3\\S_4\end{matrix} > B$	$\begin{matrix}S_3\\S_4\end{matrix} > A$	$\begin{matrix}S_3\\S_4\end{matrix} > B$	$\begin{matrix}S_2\\S_5\end{matrix} > B$	$\begin{matrix}S_3\\S_4\end{matrix} > B$	$\begin{matrix}S_9\\S_{10}\end{matrix} > B$
	$\begin{matrix}S_5\\S_6\end{matrix} > C$	$\begin{matrix}S_5\\S_6\end{matrix} > B$	$\begin{matrix}S_5\\S_6\end{matrix} > C$	$\begin{matrix}S_4\\S_6\end{matrix} > A$	$\begin{matrix}S_5\\S_6\end{matrix} > C$	$\begin{matrix}S_{11}\\S_{12}\end{matrix} > C$
Response Control	$\begin{matrix}S_1\\S_2\end{matrix} > A$	$\begin{matrix}S_1\\S_2\end{matrix} > D$	$\begin{matrix}S_1\\S_2\end{matrix} > A$	$\begin{matrix}S_1\\S_3\end{matrix} > D$	$\begin{matrix}S_1\\S_2\end{matrix} > A$	$\begin{matrix}S_7\\S_8\end{matrix} > D$
	$\begin{matrix}S_3\\S_4\end{matrix} > B$	$\begin{matrix}S_3\\S_4\end{matrix} > E$	$\begin{matrix}S_3\\S_4\end{matrix} > B$	$\begin{matrix}S_2\\S_5\end{matrix} > E$	$\begin{matrix}S_3\\S_4\end{matrix} > B$	$\begin{matrix}S_9\\S_{10}\end{matrix} > E$
	$\begin{matrix}S_5\\S_6\end{matrix} > C$	$\begin{matrix}S_5\\S_6\end{matrix} > F$	$\begin{matrix}S_5\\S_6\end{matrix} > C$	$\begin{matrix}S_4\\S_6\end{matrix} > F$	$\begin{matrix}S_5\\S_6\end{matrix} > C$	$\begin{matrix}S_{11}\\S_{12}\end{matrix} > F$

[1] The fact that the same test list was employed for all conditions is not depicted in these diagrams.

Table 2

Stimuli Employed in Mediation Experiment I

	Task I			Task II
	STIMULUS GROUPING			
	Facilitation (Intact Doublets)	Interference (Broken Doublets)	Control (New Stimuli)	
REPLICATION A				
Response Interference	Tree Gun > RL	Tree Gun > TC	Star Hat > KG	Tree Gun > KG
	Chair Bell > TC	Chair Bell > RL	Clock Fish > RL	Chair Bell > RL
	House Car > KG	House Car > KG	Cat Boat > TC	House Car > TC
Response Facilitation	Tree Gun > FW	Tree Gun > FW	Star Hat > FW	
	Chair Bell > PJ	Chair Bell > PJ	Clock Fish > PJ	
	House Car > BX	House Car > BX	Cat Boat > BX	
REPLICATION B				
Response Interference	Star Hat > PJ	Clock Fish > FW	Tree Gun > FW	Star Hat > FW
	Clock Fish > BX	Cat Boat > BX	Chair Bell > PJ	Clock Fish > PJ
	Cat Boat > FW	Star Hat > PJ	House Car > BX	Cat Boat > BX
Response Facilitation	Star Hat > KG	Clock Fish > KG	Tree Gun > RL	
	Clock Fish > RL	Cat Boat > RL	Chair Bell > TC	
	Cat Boat > TC	Star Hat > TC	House Car > KG	

Table 4

The Interaction of Response Constancy and Trials with Replication: Number Correct by Blocks of Trials, for Transfer Test

	Trials								
	1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18
Replication	Response Interference								
A	6.11	7.30	7.52	8.09	8.78	8.52	8.57	9.04	9.57
B	6.30	6.87	7.24	7.43	7.59	8.06	8.85	9.19	9.83
	Response Control								
A	6.09	6.43	7.33	7.87	7.80	8.04	8.50	9.06	9.15
B	5.46	6.85	6.57	7.48	8.29	8.26	8.94	9.00	10.07

Table 3

The Differences in Grade Levels as a Function of Replication: Trials to Criterion on List I and Number Correct on List II

List I

Grade Level Replication	First	Fifth	Retarded
A	20.25	9.08	23.25
B	16.50	7.92	17.39
Mean	17.53	13.94	15.73

List II

Grade Level Replication	First	Fifth	Retarded
A	2.47	3.31	3.19
B	2.75	3.00	2.28
Mean	2.61	3.15	2.74

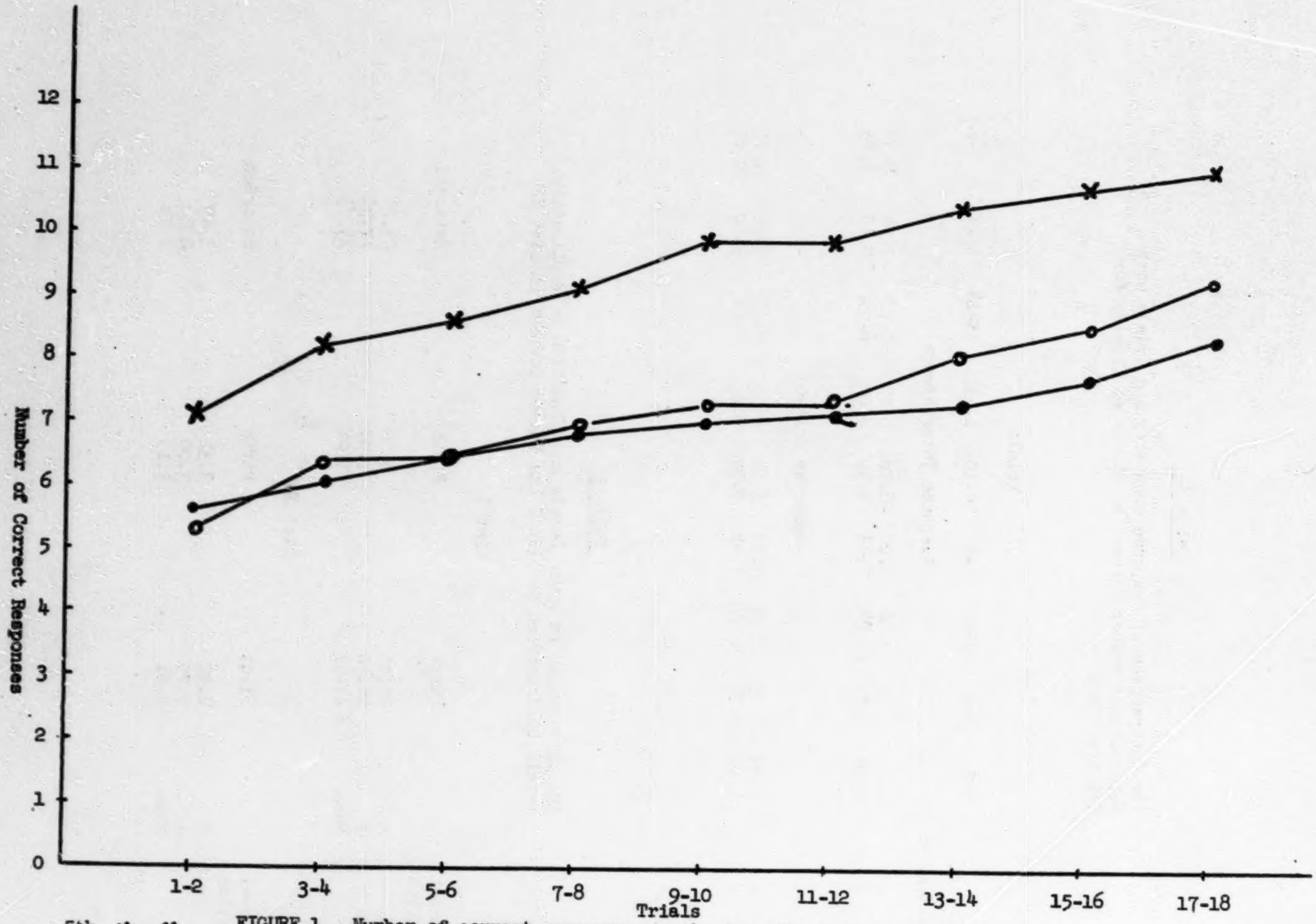


FIGURE 1. Number of correct responses on the transfer task over blocks of two trials for each group: Experiment 1.

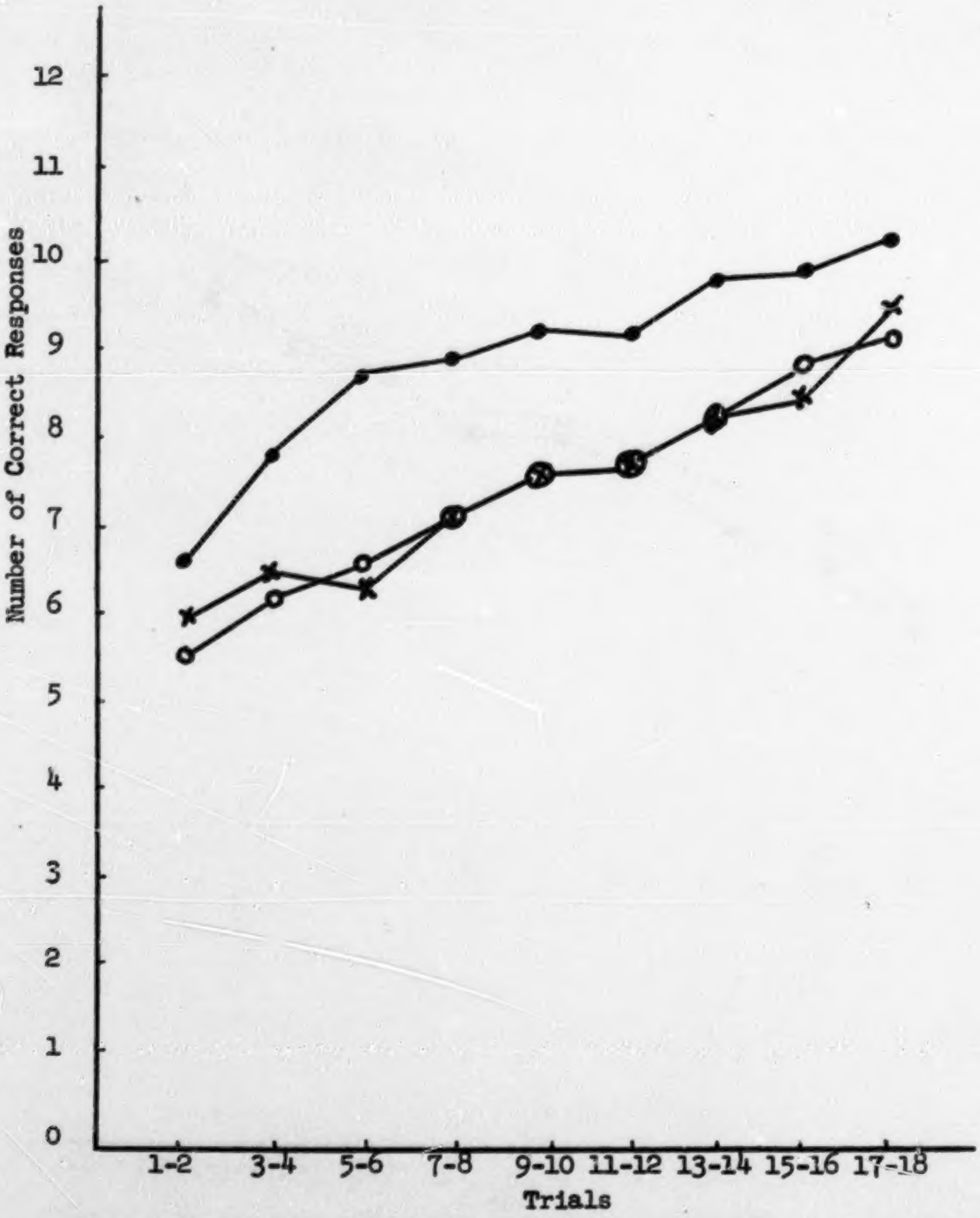


Figure 2. Number of correct responses by blocks of two trials for the three conditions of stimulus grouping: Experiment 1.

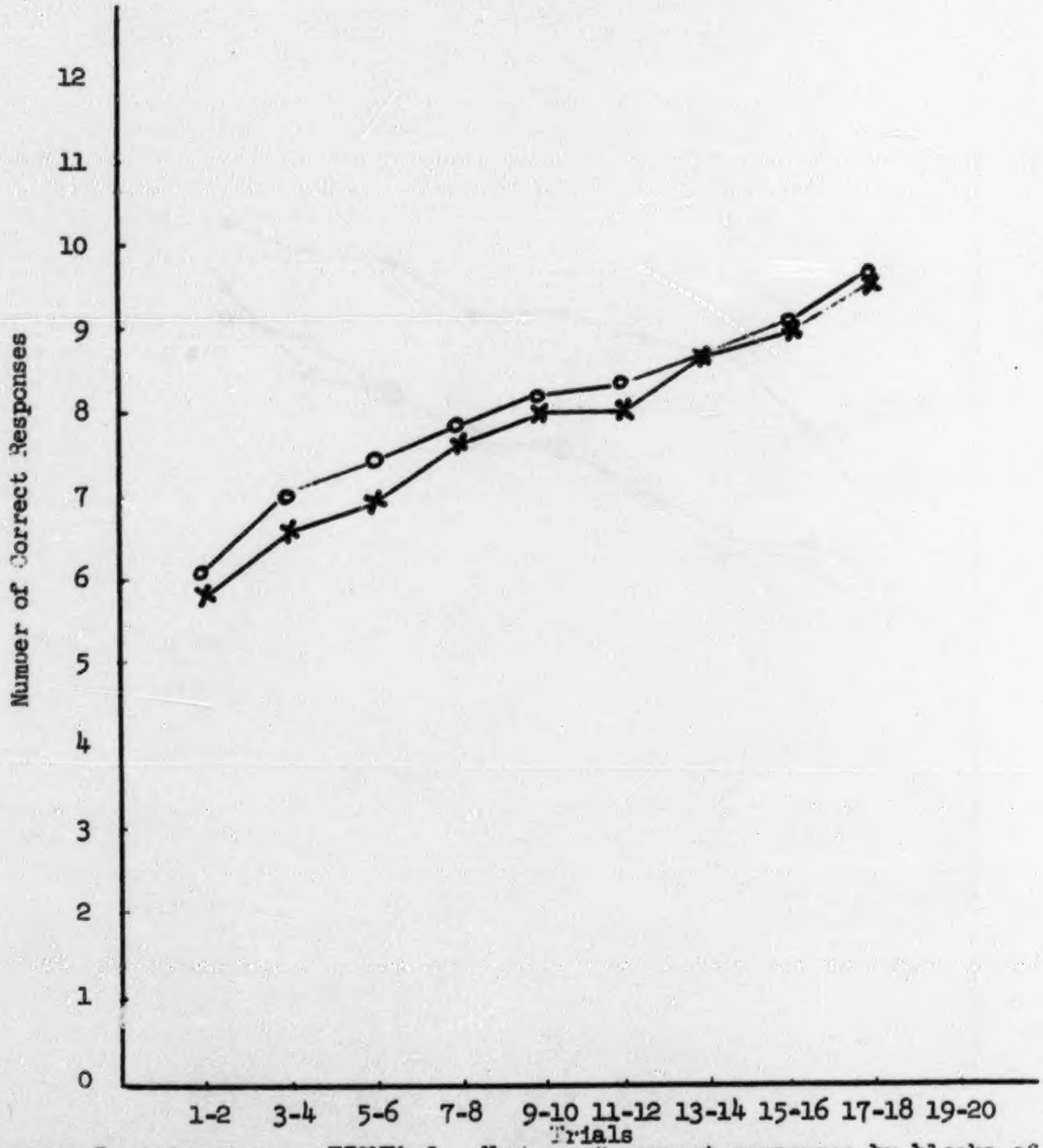


FIGURE 3. Number of correct responses by blocks of two trials for each condition of response constancy: Experiment 1.

Chapter 8

Learning Ability of Normal and Retarded Children as a Function of Set[1]

The present study is a direct outgrowth of the study described in the preceding chapter. In that study, doublets were presented during learning, on the hypothesis that fifth grade children of normal intelligence would group conceptually the stimuli belonging to the same doublet, but that first graders or retarded children might fail to do so. Analysis of the tendency to cluster during transfer suggested a dramatic departure from this hypothesis: that the older children might have deliberately suppressed clustering tendencies under the interference condition. Although as hypothesized, they did show greater clustering under the facilitation condition. Despite the plausibility of the results, there was always the nagging question of whether during original learning, all groups of children did, in fact, group the stimuli within doublets. Younger children could have grouped during learning yet not have transferred appropriately or consistently with respect to the particular transfer condition. Consequently, it was of great importance to investigate separately and independently, the effect of grouping during learning and during subsequent transfer. Thus, a set-to-group was induced in two groups of children, fifth graders and retardates, with the aid of a task which required the formation of groups during initial learning. Subsequent transfer behavior was then investigated in the present experiment (Experiment 2).

Method

Subjects

Thirty-six fifth graders and thirty-six retarded children served as Ss in Experiment 2. Intelligence test data were available for all Ss. These scores were based, for the most part, on the California Test of Mental Maturity. Fifth graders and retardates were of comparable mean chronological age, 10.8 years in both cases. But while the mean mental age of fifth graders was 100.3, that of retardates was 69.2. All Ss were attending an elementary school in the Oakland Unified School District, which draws from a low socioeconomic district.

Design and Materials

The design was a 2 x 3 x 2 factorial, with independent groups of 6 Ss per cell. The variables were Group (fifth graders and retardates), Stimulus Grouping (Facilitation, Interference, and Control), and Replication (A and B). As in the first experiment, all Ss within each replication received a formally identical task for original learning and different transfer tasks. However, as before, the specific lists employed during original learning varied in order that the test lists be identical for Ss in all conditions within each replication. The three different transfer tasks are identical with the transfer tasks employed in the first experiment. The task for original learning was changed, however. In the first experiment, during original learning, stimulus pictures within doublets were supplied with the same response, and grouping during learning was optional. For the present experiment, stimuli within doublets were directly linked by having Ss learn the stimuli as doublets directly, i.e., by presenting them as S-R and R-S pairs within a double function list. The transfer task entailed the paired-associate learning of doublets, and the transfer lists were the same as for Experiment 1. For original learning the lists contained six pairs of pictures, comprising three different pairings of six different items, each presented in the forward and the reverse order. They were presented randomly in three different orders with the restriction that no two pairs with the same members appear successively. The Stimulus Grouping conditions reflected whether for the transfer list, the stimuli which were paired during original learning were placed within the same doublet (facilitation), placed within different doublets (interference), or whether new stimuli were employed (control).

---

[1] This experiment was authored by Harriett Amster and Douglas Chalmers.

Procedure

As in Experiment 1, *E* sat next to each *S*, who was tested individually on a Stowe Memory Drum. The presentation rate throughout was again 4:2 with 6 sec. between trials. Instructions were the same as in Experiment 1, except that in Experiment 2, each *S* named all 12 pictures prior to testing to insure that all *Ss* called the pictures by the same name, and familiarization with the bigram responses was not given until the transfer task.

The procedure was also the same in the two experiments except that since only one set of three bigrams was employed for every *S* in Experiment 2, these were the only bigrams present and available during testing. *Ss* in Replication A, therefore, could refer to a card taped to the drum with KG, RL, TC listed in a column, while for Replication B a different card was used with the bigrams BX, FW, PJ. As before, the criterion for original learning was two consecutive trials for which performance was perfect.

Results

Significant differences in an analysis of variance were obtained for original learning as a function of the type of *S* ($F = 12.55$, d.f. 1/60, $p < .001$) and as a function of the particular materials employed, as illustrated by the interaction of Condition and Replication ($F = 6.38$, d.f. 2/60, $p < .01$). The fifth graders required on the average, 5.14 trials to reach a criterion of perfect responding to all six stimuli; the retardates required 10.11 trials. The measure included the first unrecorded trial but not the criterial trial. The means for the significant interaction appear in Table 2 and reflect the fact that for original learning the control list for Replication A was the same as the facilitation list for Replication B and similarly, the facilitation list for Replication B.

An analysis of performance on the first transfer trial indicated that type of *S* was the only significant main effect ($F = 8.84$, d.f. 1/60, $p < .01$) while the effect of conditions showed a trend ($F = 2.95$, d.f. 2/60, $p < .10$). The fifth graders made a mean of 3.47 correct responses on the first transfer trial while the retardates' mean was 2.39. The means for the facilitation, interference, and control conditions were, respectively, 2.38, 2.96, and 3.46. These differences could in no way be attributed to original list learning since the *F*-ratio for Conditions was > 1. The means for the interaction between Replication and Condition are presented in Table 2. Although the interaction was not significant, the *F*-ratio was 1.46, d.f. 2/60, and it is known that the difficulty of the lists employed for original learning varied among groups.

The mean number of correct responses per trial over the first eighteen transfer trials was analyzed by blocks of two trials and revealed significant main effects of type of Subject ($F = 19.52$, d.f. 1/60, $p < .005$), Condition ($F = 4.88$, d.f. 2/60, $p < .025$), and Blocks of Trials ($F = 23.12$, d.f. 8/480, $p < .005$). The means are shown in Table 3. No interactions were statistically reliable. The comparable means for Experiment 1 are included and comparison of the means for the two experiments indicates that in all cases, performance was better for Experiment 1.

For the fifth graders the tendency-to-group varied with conditions; 45% of the *Ss* showing an above-chance amount of clustering under the facilitation condition, 0% showing a comparable amount in the interference condition and 27% under control conditions. For the retardates the figures were 50%, 25%, and 8%, respectively. These values and the separate percentages for each replication are presented in Table 4. In addition, it is noteworthy that the comparable percentages for fifth graders for Experiment 1 were 33%, 0% and 33%, respectively.

Although it appeared from the means that the effect of the stimulus conditions varied as a function of the type of subjects, an analysis of variance in which the variables common to the two experiments were combined in a single analysis of variance did not show any reliable interactions. The main effects of experiment ($F = 7.44$, d.f. 1/120, $p < .01$) Grade Level ($F = 51.02$, d.f. 1/120, $p < .005$), Condition ($F = 5.81$, d.f. 2/120, $p < .005$), and Trials ($F = 4.68$, d.f. 8/960, $p < .005$) were statistically significant. The results were the same as those for the individual experiments, but showed in addition, that the performance during transfer by *Ss* in Experiment 1 was reliably superior to performance in Experiment 2. It is noteworthy that all the *F*-ratios pertaining to interactions of experimental variables with trials were > 1.

Discussion

The performance results of the experiment are highly similar to the results of Experiment 1 in that the *Ss* receiving new stimuli for transfer excelled in performance over *Ss* receiving the original stimuli arranged in the same doublets or rearranged to form new doublets. However, for the present experiment in which *Ss* were forced to form conceptual doublets, performance was poorer on the transfer list. Thus, *Ss* may be induced to learn by a conceptual process of grouping, but it is highly doubtful that grouping is facilitative and in fact, the evidence suggests that it actually impairs performance.

The fact that performance in Experiment 2 was reliably inferior to Experiment 1 provides support for the hypothesis that clustering actually interferes with learning and that learning by rote is actually more efficient in the learning of paired-associate doublets. It is of especial significance that to the extent estimated, the amount of clustering which was employed was greater for Experiment 2 than for Experiment 1 except in the case of fifth graders to whom the interference condition was administered. In this case, no *Ss* showed clustering greater than chance in either experiment. Consequently, the experimental treatment was designed to induce clustering and was found to do so in all retarded *Ss* and in fifth graders when conceivably appropriate, but the effect of this change in conceptual strategy was to uniformly hinder rather than help learning.

There are several reasons why clustering might actually hinder learning. For example, more S-R connections must be learned if clustering is adopted as a strategy than if the only connections which the *S* learns are those involved in the S and R pairs with which he is presented. In the present study, there is even a suggestion that the strategy to deliberately learn the pair within clusters as independent items may be correlated with high performance, suggesting the existence of interference from grouping, possibly deriving from generalization of errors from one *S* to the other *S* in the set. If this were true, compared to learning lists of doublets, learning should be facilitated by having *Ss* learn lists consisting of one pair from each doublet and then increasing the length of the list after learning to criterion by adding the other pair from each doublet. There is always the possibility that the present results are not conclusive for reasons other than substantive ones like those discussed above. For example, the two experiments were not conducted simultaneously by random assignment of *Ss* from the same population and thus the results are not strictly comparable. Further, by employing a different type of list during original learning, *Ss* in Experiment 2 might have suffered negative transfer when changing from picture-picture pairs to picture-bigram pairs, or having less familiarity with the responses and task requirement would have more to learn during transfer than would *Ss* in Experiment 1. In part, the learning-to-learn variable can be assessed by comparing the learning on tasks 1 and 2 in order to assess the degree of improvement under control conditions. If learning-to-learn were a dominant factor it might be supposed that the effect of conditions would increase over trials during transfer. However, there was no evidence of an interaction between conditions and trials which tends to mitigate the importance of learning-to-learn and suggests instead that the more substantive explanations be given priority.

TABLE 1

Stimuli Employed in the Second Experiment on Paired-Associate Learning in Normal and Retarded Children

REPLICATION A		REPLICATION B	
List I	List II	List I	List II
Facilitation		Facilitation	
tree - gun	tree	star - hat	star
gun - tree	> KG	hat - star	> FW
chair - bell	gun	clock - fish	hat
bell - chair		fish - clock	
house - car	chair	cat - boat	clock
car - house	> RL	boat - cat	> PJ
	bell		fish
Interference		Interference	
	house		cat
tree - chair	> TC	clock - cat	> BX
chair - tree	car	cat - clock	boat
gun - house		fish - star	
house - gun		star - fish	
bell - car		boat - hat	
car - bell		hat - boat	
Control		Control	
star - hat		tree - gun	
hat - star		gun - tree	
clock - fish		chair - bell	
fish - clock		bell - chair	
cat - boat		house - car	
boat - cat		car - house	

TABLE 2

The Mean Number of Trials to Criterion During Original Learning and Mean Number Correct on the First Transfer Trial as a Function of the List and the Replication Group

	Original Learning		Transfer	
Condition (for transfer)	Replication		Replication	
	A	B	A	B
Facilitation	5.50	10.92	3.00	1.75
Interference	7.25	6.42	2.83	3.08
Control	11.33	4.33	3.58	3.33

TABLE 3

The Main Effects of the Transfer Conditions With Respect to the Mean Number of Correct Responses per Block of Two Trials and Including Comparable Results for Experiment 1

	Experiment 1			Experiment 2		
Condition	Fifth Graders	Retardates	Mean	Fifth Graders	Retardates	Mean
Facilitation	8.77	6.61	7.69	8.62	5.66	7.14
Interference	9.41	6.44	7.92	7.80	5.44	6.62
Control	10.28	8.00	9.14	9.38	7.74	8.76
Mean	9.49	7.02		8.60	6.28	

TABLE 4

Number of Ss (%) Showing Above-Chance Grouping on the Transfer Task for Experiment 2

	NORMALS	RETARDATES
Replication A		
Facilitation	60%	67%
Interference	0%	33%
Control	33%	0%
Replication B		
Facilitation	33%	33%
Interference	0%	17%
Control	20%	17%
Combined		
Facilitation	45%	50%
Interference	0%	25%
Control	27%	8%

Chapter 9

Summary, Conclusions, and Implications

On the Effect of Variety

In spite of the vast differences among the studies in procedure and task characteristics, consistent indications appear that when large and small varieties are equal in strength, a small variety is likely to promote concept attainment to a greater extent than a large variety. In studies of mathematical concepts, the strength of the instances, i.e., the probability that each will elicit the correct concept, could be assumed to be equal for the present studies. However, in the studies of verbal concepts where the instances were unequal in strength, the mean strength of all instances within large and small variety sets of instances was held constant, and the outcomes were studied relative to the best single instance in the set. The results for convergent concepts indicated a reliable superiority of two-example sets to multiple-example sets, given that the strength of the best single instance was equal in the two cases. For the other study of the acquisition of verbal concepts, the three-example sets and six-example sets were not reliably different in the degree of approximation to the best single instance within their respective sets. But the results can be interpreted as consistent with the other experiment because of an important fact uncovered in these studies of verbal acquisition: That sets of small and large variety, equated in many respects, but not with respect to the strength of the best single instance, produced the clear outcome that the large and not the small variety produced superior concept attainment; however, control of the best single instance completely negated the superiority of the large variety.

The superiority of the large variety for associative concept attainment was demonstrated to be due to the fact that it provided the occasion for the presentation of a stronger instance than would be included in a comparable small variety. From this it may be inferred that in the absence of knowledge concerning the "goodness" of the instances of the concept-to-be-trained, a large variety should provide more correct solutions than a small variety, because of the greater probability that it will include highly probable instances. However, when instances can be assumed to be similar in strength or where the "goodness" of instances can be determined on a rational or empirical basis, a small variety of (good) instances should be optimal. It is assumed that the number of different instances presented would be logically sufficient for concept attainment.

Although conclusive evidence is not available on all possible points, it is possible to consider the role of variety in the attainment of concepts as a function of the deductive and associative processes which can be assumed to occur, albeit in varying degrees, as a function of the demands of the task and characteristics of the Ss. The experimental literature on this topic has been reviewed by the present author, partly under the auspices of this contract (Amster, 1965). In the discussion above, major consideration was given to associative determinants of concept elicitation. However, some evidence also exists with respect to other factors. For convenience, a shorthand description of the factors relevant to hypothesis testing and associative behavior are presented in the table below which also gives indication of the relative benefit of the particular factor for a small or large variety. The assumption is made that the instances vary in goodness and that large and small variety sets are equated for mean strength, but not for strength of the best single instance.

The Effect of Variety on Deductive and Associative Concept Attainment

	Small Variety	Large Variety	
1.	-	+	(Speedier rejection of false hypotheses and therefore, testing of more hypotheses per block of trials)
2.	+	-	(More false hypotheses to-be-tested) or more associations elicited which might interfere with the correct one
3.	+	-	(More false rejections of correct hypotheses because of

			the higher probability of providing very poor instances)[1]
4.	-	+	Higher probability of providing good instances
5.	+	-	Reduces memory for specific instances, hypotheses, and the outcome of their test
6.	+	-	Prominent placement of best single instance

The table illustrates certain reasons why there should be a lack of consistency in the advantage of small over large variety within and among investigators (Amster, 1965), but when hypothesis testing is unimportant (Factor 1 excluded) and the strengths of the best instances are equated for the two types of variety (Factor 4 excluded), the small variety should and where reliable differences are obtained, does excel. The factors affecting the relative effectiveness of a large and small variety in the case where the instances may be assumed to be equal in strength are the same as in the table, except that Factors 3, and 4, should become negligible.

Since the studies cover different ranges of difficulty, not to mention other sources of difference between them, the generality of the present findings may well be limited. A prime possibility for such limitation concerns the strengths of particular instances and the task difficulty. Other factors such as developmental level or type of task might well effect the relative weights of the factors as listed, or suggest the postulation of additional factors.

The factors as listed in the table reflect the operation of both deductive and associative processes. Factors 1, 2, and 3 are directly concerned with the former while Factor 1 is the most important factor concerned exclusively with deduction. Factors 2 and 3 concern both process and should effect their outcomes in similar ways. It may be inferred that in a situation where testing of hypotheses rather than elicitation of associations is a crucial determinant of effective attainment, a large variety should provide a more efficient condition unless the set includes instances which are so poor that false rejection of true hypotheses is a problem. The inclusion of poor instances is not deemed to be relevant to the associative process to a special extent since the number of false associations suggested by a poor example, may be little different from the number and strength of false associations elicited by any other example. The big difference between the good and poor examples would rest largely on the strength of the correct hypothesis.

Concept formation in a mathematical task. For the two experiments which investigated concept formation in a mathematical task, wherever there were significant effects of variety, the small variety was a more efficient condition for learning than the large variety. The characteristics of the task which may have been relevant to the superiority of the small variety consisted of the fact that instances were presented successively, which made retention a factor in attainment, and that they could be assumed equal in the probability with which they might be expected to elicit the correct response and competing responses. Analyses of the specific responses made by Ss indicated that they learned at least one of the two rules, union or intersection, and tended to repeat it on the instance which followed the correct response. Although they were never more likely to be correct in the small variety than in the large, the possibility of getting the same particular instance within the next few trials was greater in the small variety and this should have contributed to the possibility that the response would be retained. Thus, it is quite likely that retention was better in the small variety and this contributed to its advantage over the large variety.

The correct response was undoubtedly no more probable in one example than another. Thus, the large variety would not have increased the probability that a particularly "good" instance would occur but may have stimulated the elicitation of more alternative hypotheses. The pool of alternative hypotheses was likely to be similar in size from instance to instance; thus, for the small variety there were fewer possible incorrect hypotheses. These could be rejected within three trials in most small variety cases compared to rejection within one trial in the case of large variety. Thus, the inefficiency of a small variety

[1] Only relevant in certain tasks and where very low strength instances are involved.

for rejecting false hypotheses may have been a relatively unimportant factor compared with the differences in variety conditions in the number of different false hypotheses which must be rejected. An alternative was that hypothesis testing played a minor role. But if it occurs for a particular concrete situation, the relative advantage of a large over a small variety would have to be weighed in terms of the relative differences in the particular large and small varieties with respect to the need to remember previous responses and stimulus-response associations (hypotheses and the outcome of their test), the relative ease of rejecting false hypotheses, the proportion of high-probability instances in the set, and an estimate of the number of false hypotheses which would likely be generated under the two conditions. For the present study, the advantage of the small variety can be related to minimal use of hypothesis testing and emphasis on memory.

Associative concept formation in verbal tasks. The acquisition of verbal concepts involves a task which differs in major respects from the set-union task, and therefore, the fact that certain opposite effects of variety were obtained may well have been related to these differences. But is is possible to integrate the findings from these diverse tasks within a single unified framework. Among the differences most likely to be relevant to the acquisition of the concept are the fact that a) instances were presented simultaneously, rather than successively, that b) although the overall dominance of the large and small varieties were equated, individual instances were clearly unequal in the probability that they would elicit the correct response and the set included instances having high and low probabilities of eliciting this response. In this case, for acquiring the meaning of a word, the large variety of instances provided an advantage to the S in that on the average, the best instance in any large variety set had a higher probability of eliciting the response than the best instance in any small variety set despite the fact that their mean strengths were equal. Thus, it was not too surprising that the large variety proved superior to the small variety in eliciting the concept. The same data suggest that the best instance in the set is a strong determinant of its difficutly but that the strength of the poorest instance may be less important for this reason: Since a large variety not only has the most extreme best instance, it also has the most extreme worst instance on the basis of which a true hypothesis might be rejected falsely, but the results are consistent with the strength of the best instance, suggesting that hypothesis testing played an insignificant role.

Although the best instance seems a prominent determinant of difficulty, performance on the basis of the set of instances was reliably poorer than would be expected on the basis of the best single instance, since rejection of true hypotheses on the basis of poor instances should have been greater in the large variety than in the small. The fact that no difference was observed as a function of variety with respect to the superiority of the best instance over the group of instances, suggests that the benefit from good instances was more effective than the detriment from poor instances. However, the fact of the decrement relative to the best single expected score strongly suggests the operation of interference from associations to all the other instances which impede the occurrence of the correct response. This is further supported by earlier work (Musgrave, 1958; Podell, 1963C).

In short, on the basis of deductions made from the verbal acquisition experiments, a small variety would be better than a large variety if they contained the same best instance and were equal in mean probability. Thus, in deciding how to select and present instances for concept formation, it might be well to attend more to the inclusion of good instances than to the number of instances which are given. However, in the absence of information about the power of individual instances, when it is apparent that they do differ, it should be most effective to present a large variety of instances in the hope that relatively good instances would be included in the set, especially if the instances were presented simultaneously to reduce the role of memory in concept attainment.

Associative responding seemed to characterize the acquisition of verbal concepts from sentential contexts. Thus, a parallel associative situation was arranged. Convergent associations to sets of two, three, and four words having known common associations were studied as a function of the probability with which these common convergent associations had been given to the individual stimulus words. The task was formally similar to the verbal concept acquisition task in that words were presented simultaneously and were known to differ in the probability of eliciting the correct response. However, the "correct" response was in most cases far more remote than the correct response in the case of the verbal acquisition task. In any case, no consistent differences between sets of two, three, or four words was found when sets were not matched for the strength of the best single instance.

The superiority of the pairs over sets of three and four may be accounted for in part, by the fact that a good example can have more influence in a small set of instances than in a large set of instances, possibly because of its serial position (first or last) which would make it more prominent, or possibly because of the existence of less interference from competing hypotheses. The fact that the superiority of the small variety did not emerge in the study of the acquisition of word meaning may be attributable to the paucity of evidence using two instances as a small variety and three instances was found to differ little from four. However, the trend was in the same direction in both studies in that the shift occasioned by control of the best single instance was toward superiority of the small variety and away from superiority of a large variety or away from equivalence between a large and small variety.

For the sentence acquisition study results were obtained only for the easy concepts and similarly for the convergent association study the magnitude of the obtained differences increased with the strength of the best single stimulus. It would seem that the importance of the best single stimulus is increased when its strength is great. Conversely, it is likely, though not documented that the import of the worst single stimulus might be increased when its strength is very low, since this would increase the possibility of falsely rejecting a true hypothesis. In the present studies, this did not seem to be an important consideration, possibly because children were the Ss, possibly because of the tasks employed.

Conceptual grouping during learning. The paired-associate learning of doublets comprising pairs related through common convergent responses provides a kind of validation of the associative results discussed above. High variable pairs were compared with low variance pairs; the two types being equal in mean strength of eliciting the common response, but differing in the degree to which they differ. It is thus a study of the variance factor which enters into many studies of variety. It was found that for the first two learning trials fewer errors occurred for low variance doublets, but beyond that point, the high variance doublets were easier to learn. There was an overall trend (10%) for more correct responses to be made to high variance doublets, and this is consistent with the higher frequency of convergent responses which is hypothesized on the basis of the analysis of variety but was not observed. It might also be mentioned that the high and low variance doublets were equated for convergent probability which might well account for the fact that the superiority of the high variance doublets was not in evidence during the first few trials of learning. In contrast to the results with the associative task, the difference between high and low variance items decreased with the mean strength of the items. However, this might be due to a ceiling effect in which the high strength items were readily learned.

The most striking effect of variance was not in its effect on speed of learning, but in its influence upon the tendency to group items within doublets during this learning. Cluster scores and error data alike indicate that the low variance items are more likely to be learned as a group than are the high variance items. This may be due partly to the fact that there exists more occasion for such grouping when the variance is small since both items would tend to be of the same difficulty.

It is quite possible that the effects of variance on paired-associate learning are not directly related to the associative factors considered earlier. For example, the relative ease of learning high variance doublets may be due to the fact that the best single item was learned very readily, thus truncating the effective list length in which the other items were learned. In addition, if it is true that grouping during learning increases speed of learning, the tendency to group in the low variance condition might account for the greater difficulty of learning low variance doublets.

On the Effect of Mental Level

Mental ability was studied in Ss varying in age, SES, and mental age. In addition to consistent evidence of improvement in performance with age, there was evidence that the conceptual processes employed varied also. In the task involving the attainment of methematical concepts, the small variety of instances was found to benefit the low SES children, but for the high SES children there was no reliable effect of variety. This is consistent with the hypothesis that the advantage of the small variety would come about partly through its facilitation of memory, since the low SES chidlren might have poorer memories and be aided more by stimulus support than the high SES children who might tend more to memorize the relevant facts by some intentional and possibly verbal strategies. Evidence that hypothesis testing was employed to a greater extent among older children was not available on the basis of the matchematical-concept task, but the variety results are consistent with the supposition that older children and those high in SES employed hypothesis testing to a greater extent then Ss who were low in SES. Results are consistent with those of Osler and Trautman (1961).

The tendency to cluster during learning has been found to increase with age (Bousfield, Esterson, and Whitmarsh, 1958) and the present study indicates that this occurs not only for chronological age but also for mental age in that retardates cluster less than normals of comparable chronological age. This finding can be interpreted as contradictory to that of Osborn (1960). However, retardates' tendency to cluster does not differ from normals of comparable mental age. In addition, it was found that the increase in clustering with mental age is not independent of the apparent relevance of the conditions under which the clustering takes place. Normal ten-year-olds cluster more than those with a mental age of six years when clustering appears appropriate, but cluster less than others when it appears inappropriate. In fact, normal ten-year-olds cluster to an extent significantly below that which would be expected on the basis of chance, when clustering should appear inappropriate. The others, however, cluster at around the chance level regardless of the appropriateness of the conditions.

Inducing clustering in retardates does not increase their appropriate clustering more than their inappropriate clustering on a transfer test although the amount of clustering did increase. Moreover, it is possible that increasing the amount of clustering in all Ss was responsible for the obtained reduction in efficiency of learning as compared with efficiency when clustering was not induced. Thus, no evidence of increased appropriateness of clustering was occasioned by increasing the amount of clustering either in normals or retardates.

On Summation and Interference

Throughout the studies, an attempt was made to assess the possibility that the probability of occurrence of a conceptual response would be increased by the addition of stimuli which would tend to elicit it. Although such summation may well occur, it has not been observed either in the present study or in many of those conducted by others (e.g., Hill and Wickens, 1962; Musgrave and Cohen, 1964), primarily, in the present situations, because the interference effects of the additional stimuli outweigh the summative effects. Evidence for increasing interference with increasing numbers of items has been uncovered by many investigators using a wide variety of different types of tasks (e.g., Richardson, 1958; Howes and Osgood, 1954; Musgrave, 1958, 1962, Podell, 1963). However, evidence for the existence of associative priming also abounds (Cramer 1965; Mednick and Freedman, 1960; Howes and Osgood, 1954) but whether this priming entails summation must be evaluated against a baseline for assessing the predicted frequency of response from a consideration of the frequencies with which the response occurs to the individual items. Whether or not summation is reported would well depend upon the particular chance model employed. Quite a few models have been suggested (Howes and Osgood; Jenkins and Cofey, 1957; Podell, 1963) but none consistently superior.

There is some reason to speculate that the existence of summation may depend on the nature of the task and the extent to which successive instances might tend to 'narrow the sphere of meaning' rather than elicit new and conflicting spheres or hierarchies. Quite possibly, tasks which deal with the former might produce summation, while tasks which induce the latter might be more apt to reveal interference. Howes and Osgood (1954) did find four word related contexts to facilitate a type of convergent response although summation is difficult to establish Rouse and Verinis (1963) have also attempted to study this type of material. Much research on the nature of associative functioning and particularly in regard to the developmental changes in such processes remains to be done. But it is intrigueing to speculate.

Summary of Conclusions

When it is assumed that deductive reasoning is relatively unimportant, e.g. for children and for tasks in which reasoning is likely to play a small role, a small variety of instances (few different examples) is likely to produce better concept attainment than a large variety if one can assume that the instances are equally "good" (likely to elicit the correct concept).

When information about the "goodness" of instances is not available but they are believed to vary in this respect, a large variety is likely to produce better concept attainment than a small variety because it would be more likely to include "good" instances. This effect may obtain only for concepts which are fairly easy and (thus) for which "good" instances are likely to occur when the number of instances is increased.

When instances differ in "goodness", it is likely that a small variety containing the "best" instances would be superior to a large variety containing others in addition.

Repetition of the same instances in a predominantly associative task may not foster performance.

Fourth, fifth, and seventh grade children appear to rely much more heavily on associative processes than on deductive ones although a ceiling effect may have prevented the observation of deductive processes in seventh grade children in the one study which involved that age group.

Comparable results appear to occur for variations in mental ability according to grade level, SES, and mental age although the first of these seems to have the smallest effect.

The difference between an intentional and an aesthetic set on paired-associate learning was found to be more marked for fourth-grade children who were of low SES than fourth graders who were high in SES.

For children, the most important determinant of difficulty of concept formation was the strength of the best single instance in the set; yet for concept formation and convergent association the responses to the other items in the set undoubtedly interfered with concept attainment.

Grouping of items within doublets during paired-associate learning appears to hinder rather than help performance when performance within mental age levels is considered. Although under some conditions the tendency to group during training increases with mental age, performance also increases.

The learning of paired-associate doublets which are convergently related varies directly with the convergent and associative relationships between the S and R words.

A convergent response tends to be more readily attached to pairs of stimuli which have a strong "best single", i.e., in doublets having a high variance, rather than equal but highly similar in mean strength.

Conceptual grouping may come about either on the basis of the response to be trained or through another strong response.

More grouping may occur in easy and moderately easy items of equal difficulty than in difficult items because of the greater availability of mediating responses in the case of the former.

Interference from grouping may stem from the fact that more associations are learned when grouping is carried out (direct evidence of this from intrusion errors was obtained). These associations between stimuli (one of which serves as a response for the other) compete with the response-to-be-trained.

Forcing Ss to group conceptually does not enhance and may retard their performance on a transfer test. Training in rote grouping does not facilitate paired-associate learning in either normals or retardates and does not affect the appropriateness with which they use grouping in the transfer situation.

New responses are not necessarily easier to acquire than old responses when both are present and available.

Differences in learning ability as a function of mental age cannot be attributed to the tendency to form conceptual groups during learning.

Appendix 4

Materials for Experiments 2, 3, and 4 and Single Sentence Norms;

Sentences Grouped by Strict Definition of Each Concept

Finish; Complete

1. You should try to ______ the things you have left half done.
do 4; finish 18; finished 2; refinished 1; think of 1; no answer 2

2. Mary cannot ______ the problem because she doesn't understand it.
answer 4; do 7; figure 5; figure out 2; solve 3; tackle 1; think of 1; work 3; write 2

3. If his homework is hard, John doesn't ______ it.
do 22; dislike 1; finish 1; like 3; touch 1

4. Phillip asked Joan to help him ______ his homework.
correct 2; do 15; finish 5; him 1; with 2; work 1; no answer 2

5. You must have patience to ______ a job.
command 1; do 10; earn 1; finish 3; get 3; know 1; makel; take 1; tend 1; undergo 1; understand 1; wait 1; work 3

6. The painter could not ______ the room because his brush broke.
paint 27; no answer 1

Light

1. It is ______ longer in summer than in winter.
day 1; days 1; daylight 1; hot 2; hotter 3; light 3; lot 1; longer 1; much 6; night 1; notl; said 1; sunny 1; warmer 3; no answer 2

2. Large windows make a house ______.
big 5; bright 1; cool 1; dull 1; full of light 2; sunny 1; hot 1; light 4; lighter 1; nice 1; pretty 2; see through 1; ugly 1; no answer 6

3. Mother likes John to be home when it's still ______.
bright 1; dark 1; day 1; daylight 1; daytime 1; early 1; light 16; nice and warm 1; rain 1; raining 1; sunny 1; twilight 1

4. The front yard was dark, but the porch was ______.
bright 2; light 26

5. It should be ______ for reading.
bright 1; clear 1; easy 2; good 3; light 9; quiet 7; ready 1; time 1; no answer 3

6. The TV is ______.
blank 5; broke 1; broken 2; brown 1; dear 1; electrical 1; going 1; good 1; new 1; nice 1; not working 1; off 3; on 6; out 1; wrecked 1; no answer 1

Soft

1. When it got warm, the candle became ______.
bottom 1; cold 1; dim 1; dull 1; hot 6; low 1; lit 1; melted 6; melting 1; out 1; a puddle 1; smaller 2; soft 2; yellow 1; no answer 2

Continued on next page.

Soft (Cont'd)

2. If you don't want it to be heard, your whisper should be ______.
   loud 1; low 6; quiet 3; secret 1; silent 3; slow 1; soft 11; softly 1; no answer 1

3. Baby clothes are usually made in ______ colors.
   bright 6; dark 1; different 3; light 5; many 3; pink and blue 1; pretty 6; sell 1; soft 2

4. Uncle Ed likes a ______ chair after dinner.
   big 2; big soft 1; comfortable 3; comfy 1; cozy 1; easy 1; nice 1; rocking 1; soft 15; warm 2

5. A fur coat is ______ and warm.
   beautiful 3; big 3; cool 1; cozy 2; furry 3; fuzzy 2; nice 3; pretty 5; soft 5; no answer 1

6. Many people leave a ______ light in the bathroom all night.
   bathroom 1; bright 6; dim 4; good 1; lighted 1; lit 2; little 2; night 3; nightlight 1; red 1; small 2; no answer 4

Dream

1. After seeing the ghost movie, Sue hated to go to bed for fear of what she will ______.
   believe 1; cry 1; draw 1; dream 10; hear 1; imagine 1; knife 1; know 1; mine 1; nightmare 1; see 4; no answer 5

2. When Billy used to ______ he didn't know whether it was real or not.
   bring 1; dream 4; eat 1; guess 1; gun 1; imagine 2; little 2; look 2; make 3; mess 2; pass 1; read 1; snow 1; think 3; no answer 3

3. When dogs move strangely in their sleep, we wonder if they can ______.
   breathe 1; dream 12; kill 1; see 6; sleep 3; sleepwalk 2; no answer 3

4. Jimmy wants a new bicycle so badly that he will probably ______ of one.
   ask 2; brag 1; buy 7; cry 2; dream 1; faint 1; get 3; like 1; pay 2; steal 4; take 1; think 2; no answer 1

5. When George doesn't know the right answer, he may sometimes ______ one up.
   blow 1; cut it 1; crinkle 1; fail 1; give 1; make 10; pass 1; put 1; skip 2; think 6; no answer 3

6. Many people can't remember what happens when they ______.
   are born 1; crash 1; die 2; die sleep 1; dream 1; faint 1; fall 3; fall die 1; forget 3; get hurt 3; get knocked out 1; get up 1; go out 1; have amnesia 1; pass out 1; pie 1; sleep 3; no answer 2

Obstacle

1. Jane had to turn back because there was a ______ in the path.
   ball 1; black cat 1; bear 2; car 2; catamount 1; detour 1; ditch 2; dog 2; drop 1; hole 3; log 2; sign 1; skunk 1; snake 3; stream 1; tree 1; no answer 3

2. The way is clear if there is no ______ present.
   animals 1; bad 1; block 1; blockade 1; brush 1; car 1; cars 3; chalk 1; enemy 1; evil 1; fog 2; garbage 1; girls 1; man 1; nothing 1; obstruction 1; person 1; strength 1; traffic 1; trees 1; no answer 5

Continued on next page.

Obstacle (Cont'd)

3. A lazy man stops working as soon as there is a ______.
   break 2; chance 4; check 1; criminal 1; enough money 1; fight 1; fire 2; fired 1; hard job 1; job 3; rains 1; stop 1; strike 1; time 1; tired 1; weekend 1; whistle 1; no answer 4

4. A ______ keeps you from doing what you want to do.
   baby 1; belt 1; conscience 2; dad 1; feeling 1; friend 2; job 1; knife 1; love 1; mama 1; mind 1; mother 4; nerve 1; parent 5; person 1; police 1; whipping 1; no answer 2

5. Some people don't want to start a job if they know there is a ______.
   bad boss 1; bad pay 1; big man 1; bomb 1; boss 1; bully 1; catch 4; crabby boss 1; fire 1; girl 1; hard 1; hard job 1; job 1; lease 1; low pay 1; no lunch time 1; paper 2; reason 1; robber 3; time 1; trouble 1; war 1

6. He had to get rid of the ______ before finishing the task.
   bad marks 1; bees 2; blanks 1; book 1; books 1; boss 1; clothes 1; dog 1; dust 1; fly 1; guns 1; junk 1; magazine 1; marks 1; mistakes 2; paper 3; problems 1; rat 1; sentences 1; soap 1; tools 1; trash 1; water 1; words 1

Stick or Piece of Wood

1. A long ______ may be made shorter.
   board 1; cattle 1; dog 1; dog's tail 1; dress 2; foot 1; hair 1; job 1; line 1; neck 1; pencil 2; pipe 1; ribbon 2; rope 2; route 1; ruler 1; stick 1; shing 2; tail 1; word 1; no answer 3

2. A wet ______ does not burn.
   bag 1; blanket 4; board 1; cloth 1; diaper 1; fire 1; leaf 3; log 1; match 3; paper 5; rag 2; sponge 1; stop 1; water 1; no answer 2

3. You can make a ______ smooth.
   blanket 1; board 1; car 1; coat 1; dog 1; drawing 1; dress 1; face 1; fur 1; mud pie 1; paper 1; piece of wood 1; piece of tar 1; pillow 2; road 2; rock 3; rug 1; shoe 1; table 1; wood 2; word 1; no answer 2

4. The painter used a ______ to mix his paints.
   board 1; brush 3; bucket 1; finger 1; ruler 1; stick 18; stirrer 1; wooden spoon 1; no answer 1

5. Many things may be made out of a ______.
   atom 1; bed 1; cow 3; disc 1; form 1; iron 1; log 1; machine 2; molecule 1; paper 3; piece of wood 1; plank 1; rubber band 1; tree 4; wood 4; no answer 2

6. You can use a ______ for many things.
   airplane 1; bag 2; blanket 1; bike 1; book 1; box 1; hammer 2; helicopter 1; horse 1; knife 1; man 1; pencil 1; person 1; piece of wood 1; sack 1; shave 1; skirt 1; thing 1; tool 2; truck 1; wood 1; no answer 4

Happiness

1. Some people think about the things that bring them ______.
   fortunes 1; good luck 2; happiness 6; joy 3; luck 9; misfortune 1; money 1; presents 2; sad 1; toys 1; no answer 1

Continued on next page.

Happiness (Cont'd)

2. After trouble, ______ usually returns.
   bad 1; dad 1; dog 1; friends 1; fun 1; happiness 2; it 1; love 2; luck 1; more 1; people 1; policeman 3; superior 1; things 1; trouble 1; no answer 9

3. The woman remembered the ______ she had as a girl.
   background 1; book 1; doll 5; dress 1; ghost 1; life 1; man 1; movie 3; person 1; piano 1; picture 3; ring 1; sculpture 1; toys 1; tree 1; woman 2; no answer 3

4. The people at the circus usually seem full of ______.
   cheerfulness 1; drink 1; ease 1; excitement 1; food 1; fun 2; happiness 2; ice cream 1; jokes 1; joy 8; laughter 4; pep 3; talk 1; vigor 1

5. Comfort at home and a good job were all that Uncle Ben needed for ______.
   children 1; comfort 1; education 1; fun 1; happiness 3; health 1; himself 1; his family 1; his head 1; home 1; life 3; living 2; money 3; patience 1; recovery 1; rest 2; satisfaction 1; no answer 3

6. ______ and good health are found together.
   black 1; energy 1; food 3; go 1; growth 1; life 1; problems 1; rest 1; safety 1; sick 1; sickness 1; sleep 3; sports 1; strength 2; strong 1; vitamins 1; no answer 7

Stones

1. Billy collected different kinds of ______.
   coins 1; feathers 1; guns 1; insects 1; junk 1; leaves 1; models 1; money 1; newspapers 1; rocks 9; shells 2; stamps 9

2. Some ______ are found under water.
   animals 1; bugs 1; caves 2; cars 1; fish 9; fishes 1; food 2;frogs 1;mammals 1; men 1; plants 2; rocks 1; shells 4; whales 1

3. ______ often make a road very rough.
   bumps 1; dirt 1; gravel 1; holes 1; horses 1; people 1; rain 1; rock 1; rocks 14; stone 2; tractors 1; workmen 1; no answer 2

4. Only bad boys throw ______.
   beebies 1; blankets 1; eggs 1; glasses 1; mud balls 1; rock 1; rocks 17; spit balls 3; tantrums 1; blanks 1

5. A farmer does not like to find ______ in his field.
   animals 1; bugs 8; crows 4; deer 1; fox 1; grasshoppers 2; hawk 1; insects 4; weeds 3; wolves 1; no answer 2

6. ______ can polished to look smooth and shiny.
   automobiles 1; brass 1; a car 3; cars 1; copper 2; dresser 1; floor 1; guns 1; marble 2; shoe 1; shoes 10; silver 1; table 1; wood 2

Courage

1. If you have ______ you will not cry when you get hurt.
   bactine medicine 1; brave 2; bravery 1; courage 5; medicine 7; padding 1; protection 3; shots 1; no answer 7

Continued on next page.

Courage (Cont'd)

2. You need ______ to fight with a boy bigger than you.
brass knuckles 2; chalk 1; courage 6; get big 1; help 1; know how 1; muscles 4; strength 10; no answer 2

3. If you do something bad and then tell the truth, you have ______.
admit 1; admitted 1; beauty 1; been good 1; been honest 1; brave 2; a chance 1; confessed 1; courage 2; forgiven 1; honest 2; honesty 4; loyal 1; not lied 1; refined 1; told 2; trust 1; truth 2; no answer 2

4. We all admire someone who has such ______.
beauty 1; character 1; charm 2; courtesy 1; courage 1; kindness 1; knowledge 2; money 13; personality 1; sense 1; talent 1; no answer 3

5. A person who saves a baby from drowning in deep water has such ______.
brain 1; brave 1; braveness 2; bravery 2; courage 16; honor 1; love 1; power 1; shell 1; strength 1; no answer 1

6. You need ______ when you start to do a hard job.
brains 3; courage 2; education 1; energy 3; equipment 1; experience 2; food 2; help 1; muscle 1; muscles 2; noise 1; patience 4; to be smart 1; tools 1; money 1; no answer 2

Fault

1. If a suit has many ______ it is sold very cheaply.
butter 1; buttons 6; holes 10; mustaches (mistakes) 1; moths 1; nicks 1; patches 1; pockets 2; rips 2; tears 1; no answer 2

2. A person who has many ______ usually has few friends.
angers 1; bad manners 1; bully 1; dollars and toys 1; enemies 4; enemy 1; fights 3; hates 1; lacks 1; jewels 1; money 2; problem 1; troubles 1; wives 1; fault 1; no answer 7

3. People with ______ are often unhappy.
apartments 1; bad children 1; dogs 1; flu 1; hate 1; illness 1; leases 1; no children 1; polio 2; problems 1; sadness 1; sickness 1; smallpox 1; sisters 3; too many 1; trouble 2; wives 2; worm 1; no answer 5

4. Some ______ can be fixed.
automobiles 1; car 1; cars 9; chair 1; bicycles 3; bike 1; door 1; paper 1; problems 1; radios 1; things 4; tires 1; toys 2; wrecks 1; no answer 1

5. Some things are useful even if they have many ______.
bad things 1; brains 1; bugs 1; calories 1; cracks 2; disadvantages 2; faults 1; hardship 1; holes 6; knobs 1; mess 1; mistakes 1; things 1; uses 1; ways of use 1; no answer 6

6. People usually talk about the ______ of others, but not about their own.
background 1; children 4; clothes 1; face 1; fault 7; life 2; name 1; problem 3; selves 1; things 1; troubles 1; work 2; no answer 3

Appendix 4: Materials for Experiment 4

Expected Values for Each Sentence for the Strict Criterion, Consisting of the Percentage of Responses to the Single Sentences Which were Considered to be Correct Responses[1]

Sentences are Listed in Order by Concepts

1. Finish, complete, do, refinished, tackle, tend, start, fix, go back to

*1.	85.71
2.	25.00
3.	82.14
4.	71.84
5.	46.43
6.	00.00
Mean:	51.85

2. Light, bright, full of light, daytime

1.	10.71
2.	21.43
3.	60.71
*4.	100.00
5.	35.71
6.	00.00
Mean:	38.09

3. Soft, comfy, cozy

1.	7.14
2.	42.86
3.	7.14
*4.	53.57
5.	17.86
6.	00.00
Mean:	21.43

4. Dream, think, have nightmares, imagine, see, make up things, see things

1.	57.14
2.	32.14
*3.	64.28
4.	10.71
5.	21.43
6.	3.57
Mean:	31.54

---

[1] Credit given for concepts shown plus any forms of verbal or adjective and singular or plural of nouns. Underlined responses occurred in a prior study but not in experiment 4, and are not included in % of concept responses unless synonomous with given (original) concept or with others credited and occurring in verbal acquisition 4. Asterisk (*) refers to best single sentence.

5. Obstacle, block, deadend, police(man), pest, troubles, roadblock, fire, guard, rat, fly, bees, crabby boss, catamount, detour, dog, sign, skunk, snake, blockade, enemy, evil, obstruction, problems, hard job, stop, dad, mama, mother, bad boss, big man, bomb, bully, robber, war.

1.	00.00
*2.	10.71
3.	7.14
4.	3.57
5.	7.14
6.	00.00
Mean:	4.76

6. Stick or piece of wood, brush, wood, paper, match, paint brush, log, board, pipe, ruler, rock, stirrer, wooden spoon, iron, plank, box, hammer, pencil, thing, tool

1.	3.57
2.	35.71
3.	17.86
*4.	78.57
5.	32.14
6.	7.14
Mean	29.16

7. Happiness, happy, joy, pleasure, fun, gaiety, laughter, pep, luck, comfort, vigor, rest, satisfaction, strength, money, energy, playing, fortune, love, cheerfulness, ease, excitement, life, bravery

*1.	75.00
2.	14.28
3.	3.57
*4.	75.00
5.	39.28
6.	7.14
Mean:	35.71

8. Stones, rocks, shells, junk, gravel, marble

1.	39.28
2.	17.86
3.	60.71
*4.	64.28
5.	00.00
6.	00.00
Mean:	30.35

9. Courage, bravery, nerve, strength, heart, talent, know-how, honor

1.	28.57
2.	57.14
3.	14.28
4.	7.14
*5.	78.57
6.	7.14
Mean:	32.14

10. Faults, defects, mistakes, troubles, tears, wrinkles, problems, wrecks, bad things, bugs, disadvantages, mess

1.	7.14
2.	10.71
3.	10.71
4.	3.57
5.	7.14
*6.	39.28

Mean: 13.09

Mean of all Sentences = 28.81

Best single sentence mean = 65.00

*Best single sentence.

Appendix 4

The Easy and Difficult Concepts Employed for Each Variety Condition, Presented in the Order of the Frequency with Which Responses Correct by a Lenient Criterion were Given (n = 36)

	Small Variety	Small Variety Repeated	Large Variety
EASY	Stones Courage *Finish *Light Stick	Light Stones Finish *Courage *Happiness	Stones Stick Light Courage Finish
DIFFICULT	Dream Happiness Soft Fault Obstacle	Soft *Dream *Stick Obstacle Fault	Happiness *Dream *Soft Fault Obstacle

* Indicates a tie.

Total Number of Correct Responses for Each Concept Under Each Variety Condition: Lenient and (Strict) Scoring

	Faults Defects	Courage	Obstacle	Light	Stones	Happiness	Finish Complete	Soft	Stick, Piece of Wood	Dream
Small Variety (n=36)	12 (8)	21 (16)	7 (4)	20 (19)	24 (19)	15 (14)	21 (18)	14 (9)	18 (15)	16 (15)
Sm.Var. Repeated (n=36)	9 (4)	18 (15)	10 (3)	25 (25)	23 (14)	18 (15)	22 (21)	16 (10)	15 (13)	15 (12)
Large Variety (n=36)	7 (7)	24 (24)	5 (5)	24 (24)	30 (30)	19 (19)	23 (23)	13 (13)	26 (26)	18 (18)

Appendix 5

Triplets and Quartets Grouped by Part of Speech of Stimuli and the Convergent Response and Ranked in Mean Dominance (%)

TRIPLETS				QUARTETS			
Stimuli	Convergent Primary (Freq. in %)	Convergent Response (Freq. in %)	Mean Dom.%	Stimuli	Convergent Primary (Freq. in %)	Convergent Response (Freq. in %)	Mean Dom.%
Noun Stimuli - Noun Response							
Man Soldier People	Man 8	Men 3	2.46	People Lamp Kittens Spider	Animals 7	Animal 6	2.25
Foot Head Boy	Girl 22	Person 2	3.00	Cabbage Memory Hand Stomach	Food 10	Head 2	2.30
Earth Boy City	Girl 12	People 4	3.06	Sheep Comfort Dream Sickness	Dream 6 Sick 6	Bed 4	4.70
Woman Child Baby	Man 22	Mother 0	3.14	Cottage Health Salt Butter	Food 14	Food 14	5.35
People Baby Child	Man 11	Girl 3	4.34	Head Spider Hand Foot	Body 17	Leg 3	6.30
Head Shoes Hand	Body 22	Foot 6	5.20	Baby Girl Boy Child	People 22	Person 8	6.85
Spider Eagle Butterfly	Bird 12	Fly 1	6.40	Moon Bed Dream Sleep	Sun 8 Bed 8	Night 5	7.20
Priest Religion Bible	God 17	God 17	8.46	People Children Man Baby	Boy 10 People 10	Boy 10	7.20
Kittens Child Cry	Baby 14	Baby 14	8.60	Thief Child People Soldier	Man 9	Man 9	7.95

Continued on next page.

Triplets and Quartets, etc., continued

Noun Stimuli - Noun Response

Butterfly Cheese Bread	Food 20	Butter 8	9.60	Hand Health Head Stomach	Body 21	Body 21	8.50
Girl Man Woman	Boy 29	Lady 8	10.06	Man Citizen Child Children	People 14	People 14	9.00
Guns Stem Tobacco	Pipe 7	Smoke 6	12.94	Stove Stomach Fruit Cabbage	Food 15	Food 15	10.95
Ocean Cottage River	Sea 18	Lake 7	14.94	Street Window Doors Cottage	House 24	House 24	13.65
Head Foot Fingers	Body 27	Hand 8	15.80	Butterfly Dogs Sheep Lion	Cat 17	Animal 11	14.40
Boy Baby Children	Girl 21	Child 4	17.0	Man Children Woman Boy	Girl 20 People 20	Girl 20	16.65
Bed Comfort Table	Chair 18	Chair 18	17.54	Head Man People Citizen	People 8	Person 4	16.65
				Stem Salt Bath River	Water 10 Stream 10	Water 10	18.35

Noun Stimuli - Adjective Response

Lion Trouble Anger	Mad 12	Mean 3	2.20	Cabbage Anger Butter Dream	Dream 8 Sleep 8	Good 3	1.90
Chair Table Doors	House 19	Wood 1	2.34	Moon Bread Sheep Salt	Food 11	White 0	2.70

Continued on next page.

Triplets and Quartets, etc., continued
Noun Stimuli - Adjective Response

Bread Comfort Fruit	Food 20	Good 2	3.54	Mountain Ocean City House	Sea 11	Big 1	3.05
Comfort Trouble Hammer	Nail 11	Hard 5	4.06	Hand Kittens Chair Bed	Cat 7 Dog 7	Soft 2	3.25
Trouble Anger Comfort	Mad 15	Nice 2	5.06	Memory Justice Cheese Health	Doctor 7 Food 7	Good 5	6.55
Butter Carpet Comfort	Rug 8	Soft 2	5.26				
Anger Sickness Trouble	Mad 13	Bad 2	8.66				
Sickness Health Doctor	Nurse 14	Sick 7	9.34				

Adjective Stimuli - Noun Response

Bitter Salty Hungry	Food 13	Food 13	11.2	White Green Red Yellow	Color 31	Color 31	28.05

Adjective Stimuli - Adjective Response

Sweet Salty Sour	Bitter 12	Bitter 12	2.06	High Short Heavy Deep	Low 16	Big 2	2.60
High Deep Long	Low 18	Far 0	3.86	Salty Sour Bitter Sweet	Candy 11	Good 7	5.10
Swift Soft Rough	Hard 20	Smooth 7	8.60	Smooth Soft Beautiful Sweet	Hard 13	Nice 3	7.20
Salty Sweet Bitter	Sour 18	Sour 18	11.54	Yellow Soft White Heavy	Color 16	Light 6	13.20

Continued on next page.

Triplets and Quartets, etc., continued
Adjective Stimuli - Adjective Response

Yellow Green Red	Color 28	Blue 17	11.86	Red Yellow Dark White	Color 27	Black 12	16.25
Salty Bitter Sour	Sweet 18	Sweet 18	13.14	Smooth Heavy Rough Soft	Hard 23	Hard 23	18.65
Loud Deep High	Low 16	Low 16	20.8				

Appendix 5

Pairs Grouped by the Part of Speech of the Stimuli and the Convergent Response and Ranked Within Groups in Mean Dominance (%)

Stimuli	Convergent Primary (Freq. in %)	Convergent Response (Freq. in %)	Mean Dom.%
	Noun Stimuli - Noun Response		
Hand Cry	Baby 8.7	Face 3.6	.7
Earth Anger	Mad 12.7	Heaven 1.6	.7
Lamp Dream	Light 23.4	Lamb 0	.7
Wish Memory	Dream 8.3	Thought .8	.7
Carpet Window	House 11.1	Wall 1.6	.8
Bible Religion	Church 21.8	Christ .8	.8
Comfort Soldier	Army 10.3	Fort 2.8	.9
Needle Fingers	Sew 13.9	Nail .8	.9
Joy People	Happy 32.9	Laugh 3.6	1.1
Baby Whiskey	Drink 13.5	Bottle 1.6	1.9
Dogs Carpet	Cats 21.4	Pet 3.2	2.0
Man House	Lady 13.1	People .8	2.0
Children People	Man 6.7	Girl 3.6	2.1
Sheep Kittens	Cat 10.7	Dog 9.5	2.4
Blossom Stem	Flower 35.3	Tree 3.2	2.4
Cabbage Stomach	Eat 18.3	Head 2.0	2.4
Religion Cottage	House 14.7	School .8	2.9
Child Children	Baby 11.5 People 11.5	Kid 2.4	3.2
Head Girl	Boy 32.1	Hair 11.9	3.3
Justice Cold	Hot 30.2	Ice 1.6	3.3
Head Hand	Body 30.6	Leg 2.0	3.4
Cottage House	Home 20.6	Cabin 4.0	3.7
Doctor Sickness	Help 11.9	Health 10.7	4.6
Guns Stove	Fire 13.1	Fire 13.1	5.7
Priest Religion	Church 24.6	Bible 8.3	5.9
Health Wish	Good 9.9	Well 2.4	5.9
Stem Cabbage	Hot 7.1 Food 7.1	Plant 6.7	6.5
Whistle Music	Sound 10.7	Song 6.3	6.6
Earth Moon	Sun 19.0	Planet 11.9	7.8
House Cheese	Mouse 27.4	Mouse 27.4	8.3
People Child	Man 8.7	Children 7.5	10.5
Bible Music	Church 8.3	Book 5.6	11.5
Moon Light	Sun 20.2	Sun 20.2	12.3

continued on next page.

Pairs Grouped by the Part of Speech, etc., continued
Noun Stimuli - Noun Response

Woman Man	Child 21.0	Lady 8.3	12.4	Lion Dogs	Cat 24.2	Animal 8.7	18.1
Stomach Head	Body 25.8	Body 25.8	13.8	Fingers Foot	Body 18.3	Hand 16.3	21.1
Guns Tobacco	Smoke 10.7	Smoke 10.7	15.7	Table Bed	Chair 25.4	Chair 25.4	22.8
Spider Butterfly	Insect 20.6	Insect 20.6	16.2	Bread Cabbage	Food 27.8	Food 27.8	22.9
Shoes Foot	Feet 8.7	Feet 8.7	16.4	Citizen People	City 7.5	Person 2.8	24.2
Cheese Butter	Milk 17.9	Bread 9.9	17.6	Girl Joy	Boy 32.9	Boy 32.9	30.2

Noun Stimuli - Adjective Response

Memory Dream	Sleep 9.9	Thinking 0	.8	Numbers Fingers	Count 17.1	Five 7.1	4.1
Trouble Lion	Tiger 13.9	Mean 2.4	1.6	Baby Children	Child 13.9	Little 4.4	5.8
Man Boy	Girl 29.8	Male .8	1.8	Citizen Fruit	Apple 13.5	Good 6.3	8.3
Blossom Girl	Boy 25.4	Pretty 8.7	2.2	Health Sickness	Doctor 11.5	Ill 6.3	13.8
Trouble Anger	Mad 28.2	Angry 2.4	3.1	Stove Stem	Hot 23.8	Hot 23.8	18.2

Adjective Stimuli - Noun Response

Yellow Blue	Color 21.0	Pencil 0	.4	Salty Sweet	Sour 22.2	Sugar 4.0	6.2
Quiet Loud	Soft 21.8	Sound 6.0	2.5	Cold White	Snow 21.8	Snow 21.8	9.5
Dark Quiet	Light 17.9	Sleep 2.0	2.8	Hungry Bitter	Food 14.7	Food 14.7	15.1
Bitter Salty	Sour 14.7	Taste 3.2	4.7	Salty Thirsty	Water 19.4	Water 19.4	21.9
Blue High	Sky 42.5	Sky 42.5	5.2	Green Blue	Red 31.0	Color 20.2	28.1

Continued on next page.

Pairs Grouped by the Part of Speech, etc., continued

Adjective Stimuli - Adjective Response

Long Quiet	Short 19.4	Slow .8	.6
Smooth Rough	Hard 20.2	Easy .4	.6
Sweet Sour	Bitter 16.7	Bitter 16.7	2.2
Yellow Black	Color 20.2	Brown 2.4	2.3
Yellow Red	Blue 25.4	Bright .4	3.5
High Long	Short 34.5	Far .8	3.7
Red Blue	Green 28.2	Green 28.2	4.4
Long High	Short 29.8	Tall 9.5	7.0
Thinner Short	Fat 15.1	Fat 15.1	8.4
Younger Short	Tall 15.5	Small 6.0	9.4
Cold Thirsty	Hot 28.2	Hot 28.2	11.2
Red Green	Blue 27.4	Blue 27.4	14.0
Sweet Bitter	Sour 30.6	Sour 30.6	16.2
Bitter Sour	Sweet 32.5	Sweet 32.5	18.0
Short Deep	Long 20.2	Long 20.2	18.9
Black Deep	Dark 17.5	Dark 17.5	20.0
Heavy Soft	Light 27.8	Light 27.8	20.3
Beautiful White	Pretty 15.1	Pretty 15.1	21.3
Black Dark	Night 19.8	Light 18.3	21.3
Soft Rough	Hard 33.7	Hard 33.7	25.0
High Loud	Low 19.4	Low 19.4	26.6

Mixed Parts of Speech - Stimuli and Response

Bible Beautiful	God 13.1	Good 1.2	.6
Needle Slow	Fast 26.6	Poke .4	.9
Whistle Swift	Fast 11.5	Wind 3.2	1.0
Street White	Black 16.3	House 4.8	1.8
Long Doctor	Nurse 15.5 Short 15.5	Shot 2.4	1.9
Moon Blue	Sky 17.5	White 3.6	2.1
Tobacco Black	Smoke 17.1	Brown 6.3	2.2
Yellow Cabbage	Green 21.4	Green 21.4	2.3
Moon Red	Blue 11.1	Bright .4	2.4
Whistle Loud	Soft 18.3	Mouth .4	2.7
Street Deep	Hole 7.9	Long 7.1	2.8
City Heavy	Light 13.1	Big 4.4	3.3
Sour Butter	Sweet 25.8	Cream 3.6	3.6
Cottage Swift	House 17.9	Cheese 11.1	4.9

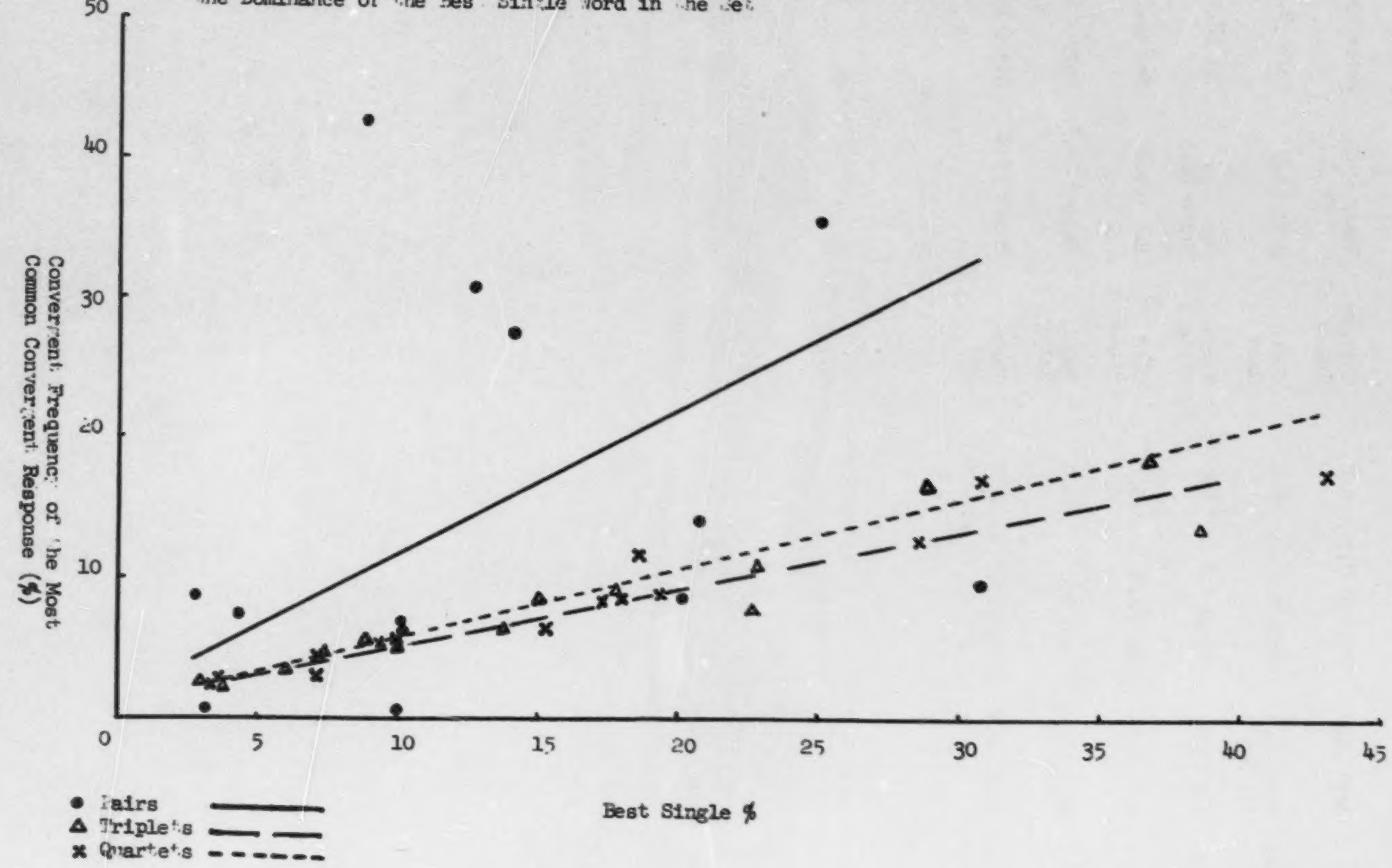
Continued on next page.

Pairs Grouped by the Part of Speech, etc., continued
Mixed Parts of Speech - Stimuli and Response

Soft Bed	Hard 19.8	Pillow 8.3	5.1
Sickness Bitter	Sick 8.3 Ill 8.3	Cold 4.0	5.1
Younger Baby	Old 13.5	Young 1.6	6.0
Salty Ocean	Sea 30.6	Salt 2.8	6.7
Sheep Light	Dark 18.3	Lamp 2.0	14.0
Deep Mountain	High 12.7	High 12.7	17.7
Thirsty Cry	Water 19.0	Water 19.0	20.5
Chair Hard	Soft 38.9	Soft 38.9	20.5
Black Kittens	Cat 18.3 White 18.3	Cat 18.3	20.7
Cold Stove	Hot 49.2	Hot 49.2	23.4
Salt Salty	Pepper 14.7	Pepper 14.7	26.7
Dark Bed	Sleep 19.4	Sleep 19.4	28.3

APPENDIX 5

The Frequency of the Most Common Convergent Response in Pairs, Triplets and Quartets Matched Item for Item in Mean Dominance and within Broader Limits for the Dominance of the Best Single Word in the Set



Appendix 6

The Four Lists Employed in Paired-Associate Learning

Convergent Strength	STRONG PRIMARY LOW VARIANCE	Mean Dominance
HIGH		
	Yellow - Green	L
	Cabbage - Green	L
	Fingers - Hand	H
	Foot - Hand	H
	Sweet - Sour	H
	Bitter - Sour	H
MEDIUM		
	People - Children	M
	Child - Children	M
	Woman - Lady	H
	Man - Lady	H
	Deep - Long	M
	Street - Long	M
LOW		
	Citizen - Good	M
	Fruit - Good	M
	Whistle - Wind	L
	Swift - Wind	L
	Window - Wall	L
	Carpet - Wall	L

Convergent Strength	STRONG PRIMARY LOW VARIANCE	Mean Dominance
HIGH		
	Lion - Animal	H
	Dogs - Animal	H
	Green - Color	H
	Blue - Color	H
	Earth - Planet	H
	Moon - Planet	H
MEDIUM		
	Quiet - Sound	L
	Loud - Sound	L
	Salty - Sugar	M
	Sweet - Sugar	M
	Priest - Bible	M
	Religion - Bible	M
LOW		
	Joy - Laugh	L
	People - Laugh	L
	Anger - Angry	L
	Trouble - Angry	L
	Butter - Cream	M
	Sour - Cream	M

	WEAK PRIMARY HIGH VARIANCE	
HIGH		
	Thinner - Fat	H
	Short - Fat	H
	Dark - Light	H
	Black - Light	H
	Deep - High	H
	Mountain - High	H
MEDIUM		
	Cabbage - Plant	M
	Stem - Plant	M
	Doctor - Health	M
	Sickness - Health	M
	Kittens - Dog	M
	Sheep - Dog	M
LOW		
	Blue - White	L
	Moon - White	L
	Baby - Bottle	L
	Whiskey - Bottle	L
	Hand - Face	L
	Cry - Face	L

	STRONG PRIMARY HIGH VARIANCE	
HIGH		
	Bed - Sleep	H
	Dark - Sleep	H
	Stove - Hot	H
	Stem - Hot	H
	Kittens - Cat	H
	Black - Cat	H
MEDIUM		
	Women - Men	M
	Thief - Men	M
	Girl - Hair	L
	Head - Hair	L
	Cheese - Bread	M
	Butter - Bread	M
LOW		
	Quiet - Slow	L
	Long - Slow	L
	Salty - Salt	M
	Ocean - Salt	M
	House - Cabin	L
	Cottage - Cabin	L

Experiment I

Appendix 7

TABLE 1

Distribution of Ss in Schools by Grade and Replication. Number of Different Classes (and Teachers) are Given in Parentheses.

School	First		Fifth		Retardates	
	A	B	A	B	A	B
1						36
2		12(1)		12(1)	27	
3		24(3)		24(2)	7	
4	36(3)		36(3)		2	
Total	36	36	36	36	36	36

TABLE 2

Means and Ranges of CA and IQ as a Function of Grade and Replication. Number of Ss Upon Which the Means are Based are Given in Parentheses.

	First		Fifth		Retardates	
	A	B	A	B	A	B
C.A.	6.37(35)	6.51(35)	10.36(36)	10.53(36)	10.56(36)	10.64(36)
Mean	6.44(70)		10.44(72)		10.60(72)	
Range	6-7		10-11		9-13	
I.Q.			108.23(35)	110.94(35)	69.21(24)	68.92(36)
Mean	100 (Assumed)		109.59(70)		69.03(60)	
Range			78-141		48-80	
M.A.						
Mean	6.44 (Assumed)		11.45		7.32	
Range						

Appendix 7

Subject Selection[1]

All Ss were enrolled in one of the San Leandro, California elementary schools: Washington, Cleveland, Madison, or Halcyon. The distribution of schools for the different groups and replications are shown in Table 1. The identity of the schools is available from the authors with proper justification. Ss were drawn from one class at a time, alphabetically, until the supply was depleted. Miss Alice Gordon ran the first two Ss in each of the six conditions of Phase A for first and fifth graders. Miss Barbara Juster ran all the remaining 192 Ss.

The mean chronological and mental ages of the Ss are given in appendix 1, and 2. It was possible to secure the C. A.s of all Ss except two of the first graders. All C. A.s were computed from knowledge of each S's age in years from his last birthday.

IQs were not available for first graders; so it was assumed that their mean M. A.s approximately matched their mean M. A. IQs were, however, available for all but two of the normal fifth graders.

All but nine of the retarded children (i.e., all those from Cleveland and Washington schools) were selected from special education classes, for children with a tested IQ of less than 75. All retardates used in this experiment scored below the .05 percentile on the Reading Vocabulary Section of the 1957 short form of the CAT (elementary forms W and Y). The IQs given in Table 2 for both retardates and normals are based upon the highest score for each S on individual tests which were either the CTMM, S - B, or the Wechsler-Bellevue. This information was available for only 60 of the retarded Ss.

---

[1] Prepared by D. Chalmers.

Appendix 8

Subject Characteristics

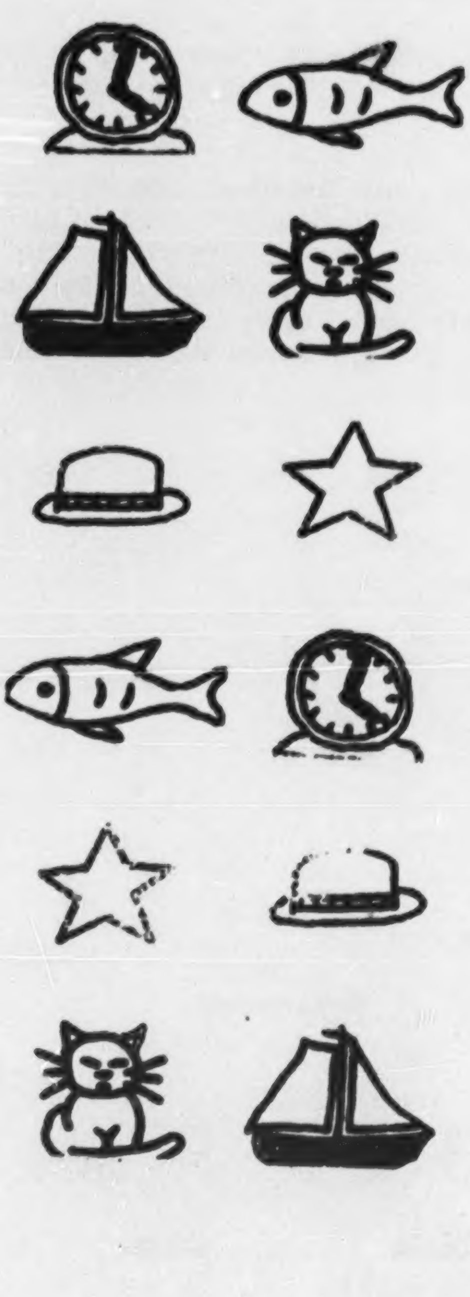
	Fifth Graders		Retardates	
	A	B	A	B
Age	10.94 (N = 18)	10.56 (N = 16)	10.63 (N = 16)	10.89 (N = 18)
I.Q.	96.37 (N = 16)	104.23 (N = 13)	68.31 (N = 16)	70.11 (N = 18)

All Ss were attending Lockwood Elementary School in Oakland. Fifth graders were attending the summer session and were tested in July of 1965. Retardates were in special classes during the regular session and were tested in June, 1965.

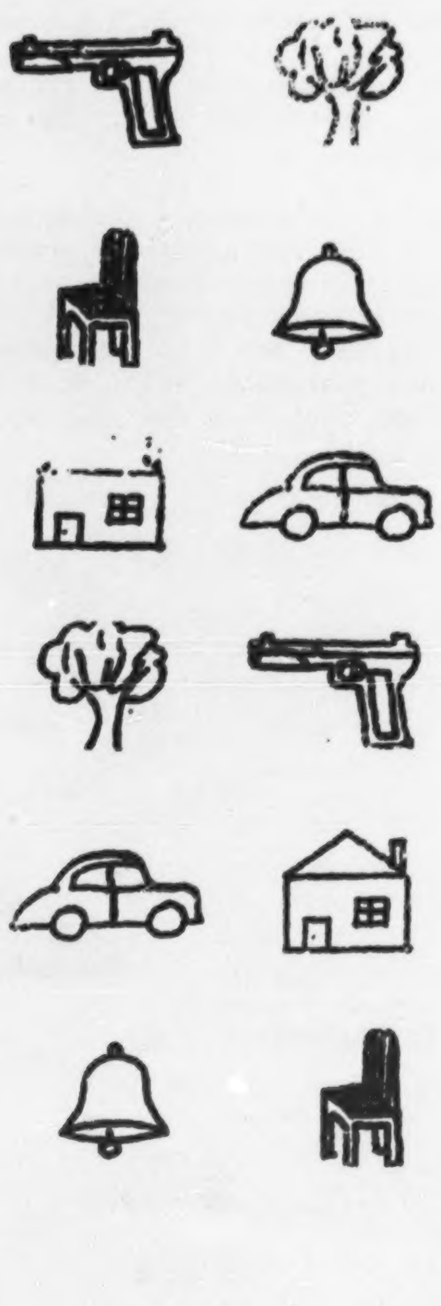
Appendix 8

Pictorial Stimuli as Paired in Experiment 2

A



B



REFERENCES

Adams, J. A. Multiple vs. single problem training in human problem solving. *J. exp. Psychol.*, 1954, 48, 15.

Amster, H. The effect of the variety of instances on the production of verbal concepts. Am. Psychologist, 1963, 18, 382 (abstract).

*Amster, H. The effect of instructional set and variety of instances in children's learning, 1964. In press, *J. educ. Psychol.*, 1966.

*Amster, H. concept formation in children, *Elem. Eng*. 1965, XXXXII, 543-552.

*Amster, H. Association as a process in the acquisition of semantic meaning from sentential contexts. Paper presented at Western Psych. Assoc., 1965.

Amster, H. and Keppel, G. Convergent association norms for children and adults. To appear in *Psychonomic Science Monographs*, 1966.

Amster, H. and Keppel, G. A developmental study of the acquisition and retention of letter pairs which vary in associative strength. Paper presented at the annual meeting of the Psychonomic Society, 1965.

*Amster, H. and Marascuilo, L. Effect of type of pretraining and variety of instances on children's concept learning. *J. exp. child Psychol.*, 1965, 2, 192-204.

Attneave, R. and Arnoult, M. D. The quantitative study of shape and pattern perception. *Psychol. Bull.*, 1956, 53, 452-471.

Bousfield, W. A., Esterson, S. and Whitmarsh, G. A. A study of developmental changes in conceptual and perceptual associative clustering. *J. genet. Psychol.*, 1958, 92, 95-102.

Bruner, J. S., Goodnow, J. J. and Austin, G. A. *A Study of Thinking*. New: Wiley, 1956.

Callantine, M. F. and Warren, J. M. Learning sets in human concept formation. *Psych. Rep.*, 1955, I, 363-367.

Carlin, J. Word-association strength as a variable in verbal paired-associate learning. 1958 Ph. D. thesis, U. of Minnesota.

Cramer, P. H. Mediated priming of associative responses: the effect of time lapse and interpolated activity. Presented at E. P. A., 1965.

Ervin, S. M. Changes with age in the verbal determinants of word-association. *Am. J. Psychol.*, 1961, 74, 361.

Fields, P. E. Studies in concept formation: I. The development of the concept of triangularity in the white rat. *Comp. Psychol. Monogr.*, 1932, 9, No. 42.

Gagné, R. M. and Bassler, O. C. Study of retention of some topics of elementary non-metric geometry. *J. educ. Psychol.*, 1963, 54, 123-131.

Gagné, R. M., Mayor, J. R., Garstens, H. R. and Paradise, N. E. Factors in acquiring knowledge of a mathematical task. *Psychol. Monogr.*, 1962, 76 (7, Whole No. 526).

Harlow, H. F. Thinking. In H. Helson (Ed.), *Theoretical Foundations of Psychology*. New York: Van Nostrand, 1951.

Hill, F. A. and Wickens, D. D. The effect of stimulus compounding in paired-associate learning. JvLvB, 1, 144-151, 1962.

Hovland, C. I. and Morrisett, L. A comparison of three varieties of training in human problem solving. *J. Exp. Psychol.*, 1959, Vol. 58, No. 1, 52-54.

References, continued.

Howes, D. and Osgood, C. E. On the combination of associative probabilities in linguistic contexts. Amer. J. Psychol., 1954, 67, 241-258.

Inhelder, Barbel and Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Jenkins, P. M. and Cofer, C. N. An exploratory study of discrete free association to compound verbal stimuli. Psych. Rep. 1957, 3, 599-602.

Kendler, H. H. and Kendler, T. Vertical and horizontal processes in problem solving. Psych. Rev., 1, 1962.

*Marascuilo,L. and Amster, H. The effect of variety in children's concept learning. Calif. J. educ. Res., 1966, In press.

Mednick, S. A. and Freedman, J. L. Facilitation of concept formation through mediated generalization. J. exp. Psychol., 1960, 60, 278-283.

Musgrave, B. S. Context effects on word associations using one-word, two-word and three word stimuli. Paper presented at E. P. A., 1958.

Musgrave, B. S. The effect of nonsense-syllable compound stimuli on latency in a verbal paired-associate task. J. exp. Psychol., 1962, 63, 499-504.

Musgrave, B. S. and Cohen, J. C. Effects of two-word stimuli on recall and learning in a paired-associate task. J. exp. Psychol., 1964, 68, 161-166.

Osborn, W. J. Associative clustering in organic and familiar retardates. Amer. j. ment. Devic., 1960, 65, 351-357.

Osler, S. F. and Fivel, M. W. Concept attainment: I. Effect of age and intelligence in concept attainment by induction. J. exp. Psychol., 1961, 62, 1-8.

Osler, S. F. and Trautman, G. E. Concept attainment: II. Effect of stimulus complexity upon concept attainment at two levels of intelligence. J. exp. Psychol., 1961, 1, 9-13.

Palermo, D. S. Mediated association in a paired-associate transfer task. J. exp. Psychol., 1962, 3, 234-238.

Palermo, D. S. Word associations and children's verbal behavior. I. In L. P. Lipsitt and C. C. Spiker (Eds.). Advances in child development and behavior. New York: Academic Press, 1963.

Palermo, D. S. and Jenkins, J. J. Paired-associate learning as a function of the strength of links in the associative chain, 1965, J. verb. Learn. verb. Behav., 3, 406-412.

Palermo, D. S. and Jenkins, J. J. Word association norms: Grade school through college. Minneapolis: U. of Minnesota Press, 1964.

Podell, H. A. Two processes of concept formation. Psychol. Monogr., 1958, 72.

Podell, H. A. The effect of the variety of instances on the production of verbal concepts. Amer. Psychologist, 1963, 18, 382 (Abstract) (a).

Podell, H. A. A developmental study of recall of instances as a function of set and the variety of the instances presented. Amer. Psychologist, 1963, 18, 346. (Abstract) (b).

Podell, H. A. A quantitative study of convergent association. J. verb. Learn. verb. Behav., 1963, 2, 234-241. (c).

References, continued.

Rasmussen, E. A. and Archer, E. J. Concept identification as a function of language pretraining and task complexity. J. exp. Psychol., 1961, 61, 437-441.

Richardson, J. The relationship of stimulus similarity and number of responses. J. exp. Psychol., 1958, 56, 478-484.

Richardson, J. Association among stimuli and the learning of verbal concept lists. J. exp. Psychol., 1960, 5, 290-298.

Richardson, J. The learning of concept names mediated by concept examples. J. verb. Learn. verb. Behav., 1963, 1, 281-288.

Rouse, R. O. and Verinis, J. S. The effect of compound verbal stimuli on word association. Paper presented at E. P. A., 1963.

Shapiro, S. S. Paired associates learning in children: Length of list, initial associative strength, presentation time and grade level. Manuscript, 1965.

Tulving, E. Subjective organization in the free recall of "unrelated" words. Psychol. Rev., 1962, 69, 344-354.

Vinacke, W. E. Concept formation in children of school ages. Education, 1954, 74, 527-534.

Werner, H. and Kaplan, E. The acquisition of word meaning. Child Development Monogr., 1951.

Whorf, B. L. Language, Thought and Reality. New York: Wiley, 1956.

Wicklund, D. A. Palermo, D. S. and Jenkins, J. J. The effects of associative strength and response hierarchy on paired-associate learning. J. verb. Learn. verb. Behav., 1964, 3, 413-420.

Wicklund, D. A., Palermo, D. S. and Jenkins, J. J. Associative clustering in the recall of children as a function of verbal association strength. J. exp. child Psychol., 1965, 2, 58-66.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

* Papers prepared under the auspices of this contract.